

A CONTINUOUS MODEL FOR SYSTEMS OF COMPLEXITY 2 ON SIMPLE ABELIAN GROUPS

PABLO CANDELA AND BALÁZS SZEGEDY

ABSTRACT. It is known that if p is a sufficiently large prime then for every function $f : \mathbb{Z}_p \rightarrow [0, 1]$ there exists a continuous function on the circle $f' : \mathbb{T} \rightarrow [0, 1]$ such that the averages of f and f' across any prescribed system of linear forms of complexity 1 differ by at most ϵ . This result follows from work of Sisask, building on Fourier-analytic arguments of Croot that answered a question of Green. We generalize this result to systems of complexity at most 2, replacing \mathbb{T} with the torus \mathbb{T}^2 equipped with a specific filtration. To this end we use a notion of modelling for filtered nilmanifolds, that we define in terms of equidistributed maps, and we combine this with tools of quadratic Fourier analysis. Our results yield expressions on the torus for limits of combinatorial quantities involving systems of complexity 2 on \mathbb{Z}_p . For instance, let $m_4(\alpha, \mathbb{Z}_p)$ denote the minimum, over all sets $A \subseteq \mathbb{Z}_p$ of cardinality at least αp , of the density of 4-term arithmetic progressions inside A . We show that $\lim_{p \rightarrow \infty} m_4(\alpha, \mathbb{Z}_p)$ is equal to the infimum, over all continuous functions $f : \mathbb{T}^2 \rightarrow [0, 1]$ with $\int_{\mathbb{T}^2} f \geq \alpha$, of the following integral:

$$\int_{\mathbb{T}^5} f\left(\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right) f\left(\begin{smallmatrix} x_1 + x_2 \\ y_1 + y_2 \end{smallmatrix}\right) f\left(\begin{smallmatrix} x_1 + 2x_2 \\ y_1 + 2y_2 + y_3 \end{smallmatrix}\right) f\left(\begin{smallmatrix} x_1 + 3x_2 \\ y_1 + 3y_2 + 3y_3 \end{smallmatrix}\right) d\mu_{\mathbb{T}^5}(x_1, x_2, y_1, y_2, y_3).$$

1. INTRODUCTION

One of the well-known central objectives in arithmetic combinatorics is to find optimal bounds for Szemerédi's theorem on arithmetic progressions. A closely related problem is to determine how small the average across k -term progressions can be for a $[0, 1]$ -valued function of fixed average value on a large cyclic group $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$. We write this quantity as follows:¹

$$m_k(\alpha, \mathbb{Z}_N) := \inf_{f : \mathbb{Z}_N \rightarrow [0, 1], \mathbb{E}_{\mathbb{Z}_N} f \geq \alpha} \mathbb{E}_{n_1, n_2 \in \mathbb{Z}_N} f(n_1) f(n_1 + n_2) \cdots f(n_1 + (k-1)n_2).$$

A natural direction in which to gain insight on this difficult problem consists in analysing the asymptotic behaviour of these quantities as $N \rightarrow \infty$. Answering a question of Green

2010 *Mathematics Subject Classification.* Primary 11B30; Secondary 11K70, 43A85.

Key words and phrases. Linear forms, higher-order Fourier analysis, nilmanifolds, equidistribution.

¹For a finite set X and a function $f : X \rightarrow \mathbb{C}$, we denote by $\mathbb{E}_{x \in X} f(x)$ the average $\frac{1}{|X|} \sum_{x \in X} f(x)$.

[4, Problem 3.1], Croot took a first step in this direction for the case $k = 3$, by proving the following result [3, Theorem 1].

Theorem 1.1 (Croot). *For every $\alpha \in [0, 1]$, the sequence $m_3(\alpha, \mathbb{Z}_N)$ converges as $N \rightarrow \infty$ through the primes.*

The main tool in Croot's proof is the Fourier transform on \mathbb{Z}_N . More recently, using higher-order Fourier analysis, the first named author and Sisask extended this convergence result to k -term arithmetic progressions for every k , and more generally to all systems of linear forms of finite complexity² [2, Theorem 1.3]. Thus, provided the values of N are adequately restricted, the problem of estimating $m_k(\alpha, \mathbb{Z}_N)$ has a well-defined asymptotic version, namely to estimate the following limit as a function of α :

$$\lim_{\substack{N \rightarrow \infty \\ N \text{ prime}}} m_k(\alpha, \mathbb{Z}_N), \quad (1)$$

and especially to determine the order of magnitude of this function as $\alpha \rightarrow 0$.

To shed light on this problem, it is natural to seek an expression for the limit in terms of an integral over some fixed object, hopefully some compact abelian group as simple as possible. A result in this direction was given for $k = 3$ by Sisask, identifying the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ as such a group, as follows [20, Theorem 4.1.2].

Theorem 1.2 (Sisask). *For every $\alpha \in [0, 1]$, we have*

$$\lim_{p \rightarrow \infty} m_3(\alpha, \mathbb{Z}_p) = \inf_{\substack{f: \mathbb{T} \rightarrow [0, 1] \text{ Borel} \\ \int_{\mathbb{T}} f \geq \alpha}} \int_{\mathbb{T}^2} f(x_1) f(x_1 + x_2) f(x_1 + 2x_2) \, d\mu_{\mathbb{T}^2}(x_1, x_2).$$

(Here and below, given a compact abelian group Z we denote by μ_Z the Haar probability measure on Z , and p denotes a prime number.)

We are thus led to the question of what would be an adequate generalization of Theorem 1.2 for k -term progressions with $k > 3$. The main result of this paper provides an answer for $k = 4$. In order to state our result let us gather some terminology.

Definition 1.3 (Ψ -average over a compact abelian group). Let D, t be positive integers and let $\Psi = (\psi_1, \dots, \psi_t)$ be a system of integer linear forms, thus $\psi_i : \mathbb{Z}^D \rightarrow \mathbb{Z}$ is a homomorphism for each $i \in [t]$. For a compact abelian group Z and a measurable

²We recall this notion of complexity in Definition A.7. Let us recall also that the infimum in $m_k(\alpha, \mathbb{Z}_N)$ can be restricted to indicator functions of subsets of \mathbb{Z}_N without affecting its asymptotic behaviour, as explained in [3, p. 3] for $k = 3$; for $k > 3$ one can argue similarly, using [2, Lemma 8.2].

function $f : Z \rightarrow \mathbb{C}$, we define³

$$S_\Psi(f : Z) = \int_{Z^D} f(\psi_1(x)) \cdots f(\psi_t(x)) \, d\mu_{Z^D}(x). \quad (2)$$

We call this the Ψ -average of f (or average of f across Ψ) over Z .

We need an analogous notion of averaging across systems of linear forms for functions on a *filtered nilmanifold* $X = (G/\Gamma, G_\bullet)$. This analogue relies on the concept of the *Leibman nilmanifold* associated with a system $\Psi : \mathbb{Z}^D \rightarrow \mathbb{Z}^t$ and X . This is a certain subnilmanifold of X^t that we shall denote by X^Ψ . These concepts have been used in previous works (see [13]) and we recall their definitions in Appendix A. For now let us take these concepts for granted to give the following main definition.

Definition 1.4 (Ψ -average over a filtered nilmanifold). Let D, t be positive integers and let $\Psi = (\psi_1, \dots, \psi_t)$ be a system of linear forms $\psi_1, \dots, \psi_t : \mathbb{Z}^D \rightarrow \mathbb{Z}$. Let $X = (G/\Gamma, G_\bullet)$ be a filtered nilmanifold (see Definition A.4), let X^Ψ be the Leibman nilmanifold for Ψ on X , and let $f : G/\Gamma \rightarrow \mathbb{C}$ be measurable.⁴ We define the Ψ -average of f over X to be

$$S_\Psi(f : X) = \int_{X^\Psi} f(y_1) \cdots f(y_t) \, d\mu_{X^\Psi}(y). \quad (3)$$

We define $m_\Psi(\alpha, X) = \inf \{ S_\Psi(f : X) \mid f : X \rightarrow [0, 1] \text{ Borel, } \int_X f \, d\mu_X \geq \alpha \}$.⁵

As a first example, one can check that letting X denote \mathbb{T} equipped with the lower central series (this is a nilmanifold that happens to be an abelian group), then, using Definitions A.1 and A.6, one has that $S_\Psi(f : X)$ is just the average $S_\Psi(f : \mathbb{T})$ in (2).

An important aspect of Definition 1.4 is that the choice of filtration in X can modify the Leibman nilmanifold X^Ψ significantly and thereby change the Ψ -average (we illustrate this in Example 1.6 below).

From now on, given a group G , we shall denote by $G_{\bullet(1)}$ the lower central series on G . Denoting by X_1 the filtered nilmanifold $(\mathbb{T}, \mathbb{R}_{\bullet(1)})$, Theorem 1.2 can be rephrased as saying that $\lim_{p \rightarrow \infty} m_{\Psi_3}(\alpha, \mathbb{Z}_p) = m_{\Psi_3}(\alpha, X_1)$, where Ψ_3 is the system corresponding to 3-term arithmetic progressions. It is known that this result holds more generally in that Ψ_3 may be replaced with any system of complexity 1. The point of our main result is that by replacing X_1 with a natural and only slightly more complicated filtered group,

³Each map ψ_i is originally defined on \mathbb{Z}^D but the definition then extends to any Z^D . Thus on \mathbb{Z}^D we have $\psi_i(n) = c_{i,1}n_1 + \cdots + c_{i,D}n_D$ for some $c_{i,j} \in \mathbb{Z}$, and then $\psi_i(x) = c_{i,1}x_1 + \cdots + c_{i,D}x_D$ on Z^D .

⁴Measurability in this paper is always with respect to the Borel σ -algebra. Recall also that every nilmanifold $X = G/\Gamma$ has a G -invariant Borel probability measure, which we denote by μ_X .

⁵If each coordinate projection $X^t \rightarrow X$ restricts to a measure preserving map from X^Ψ to X , then by a simple argument using Lusin's theorem (discussed in Remark 2.9 below) one sees that the value of $m_\Psi(\alpha, X)$ is unchanged if we restrict the infimum to continuous functions $f : X \rightarrow [0, 1]$.

one obtains an extension of Theorem 1.2 for all systems of complexity at most 2.

If G_\bullet, H_\bullet are filtrations on groups G, H , with i -th terms $G_{(i)}, H_{(i)}$, then the *product filtration*, denoted $G_\bullet \times H_\bullet$, is the filtration on the direct product $G \times H$ with i -th term $G_{(i)} \times H_{(i)}$. For an integer $d \geq 2$ and an abelian group Z , we denote by $Z_{\bullet(d)}$ the *maximal degree- d filtration* on Z , that is the filtration with i -th term equal to Z for $i = 0, \dots, d$ and equal to $\{0_Z\}$ for $i > d$. We can now state our main result.

Theorem 1.5. *Let $X_2 = (\mathbb{T}^2, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)})$. Then for every system Ψ of integer linear forms of complexity at most 2, and every $\alpha \in [0, 1]$, we have*

$$\lim_{p \rightarrow \infty} m_\Psi(\alpha, \mathbb{Z}_p) = m_\Psi(\alpha, X_2). \quad (4)$$

To illustrate both the result and some of the underlying notions, let us pause to give a more explicit description of (4) in a specific case.

Example 1.6. Consider the system $\Psi_4 : n \in \mathbb{Z}^2 \mapsto (n_1, n_1 + n_2, n_1 + 2n_2, n_1 + 3n_2)$ corresponding to 4-term arithmetic progressions, a central example of a system of complexity 2. On X_1 , the Leibman nilmanifold corresponding to Ψ_4 is the ‘usual’ 2-dimensional subgroup of \mathbb{T}^4 consisting of all 4-term progressions: $X_1^{\Psi_4} = \{(x_1, x_1 + x_2, x_1 + 2x_2, x_1 + 3x_2) : x_1, x_2 \in \mathbb{T}\}$. If instead we equip \mathbb{T} with the filtration $\mathbb{R}_{\bullet(2)}$, letting $Y = (\mathbb{T}, \mathbb{R}_{\bullet(2)})$, then the corresponding Leibman nilmanifold for Ψ_4 is the following 3-dimensional subgroup of \mathbb{T}^4 , which contains $X_1^{\Psi_4}$: $Y^{\Psi_4} = \{(y_1, y_1 + y_2, y_1 + 2y_2 + y_3, y_1 + 3y_2 + 3y_3) : y_i \in \mathbb{T}\}$. (Similar calculations have been used as examples in previous works; see in particular [13, Section 3].)

Now, on X_2 , the Leibman nilmanifold for Ψ_4 is $X_1^{\Psi_4} \times Y^{\Psi_4}$, a 5-dimensional subtorus of X_2^4 . Then, Theorem 1.5 tells us that $\lim_{p \rightarrow \infty} m_{\Psi_4}(\alpha, \mathbb{Z}_p)$ is equal to the infimum, taken over all continuous functions $f : \mathbb{T}^2 \rightarrow [0, 1]$ with $\int_{\mathbb{T}^2} f \geq \alpha$, of the following integral:

$$\int_{\mathbb{T}^5} f \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} f \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} f \begin{pmatrix} x_1 + 2x_2 \\ y_1 + 2y_2 + y_3 \end{pmatrix} f \begin{pmatrix} x_1 + 3x_2 \\ y_1 + 3y_2 + 3y_3 \end{pmatrix} d\mu_{\mathbb{T}^5}(x_1, x_2, y_1, y_2, y_3),$$

where we write elements θ of \mathbb{T}^2 in column form, $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$.

Our proof of Theorem 1.5 relies on a *transference result* for systems of complexity 2, stated in Section 2 (see Theorem 2.1). Roughly speaking, the main part of this result tells us that given a $[0, 1]$ -valued function f_1 on a large group \mathbb{Z}_p , there is a continuous $[0, 1]$ -valued function f_2 on X_2 that resembles f_1 in the sense that certain averages of f_2 across systems of complexity at most 2 have roughly the same value as those of f_1 . To motivate our choice of X_2 further, in Subsection 2.1 we show that, if f_1, f_2 are allowed to be complex-valued, then such a result does not hold if instead of X_2 we use X_1 .

The proof of our transference result occupies most of the rest of the paper. It relies on the regularity method for the Gowers U^3 norm, and also on a notion of *modelling* between filtered nilmanifolds that is central to this paper, defined in Subsection 2.2 (see Definitions 2.11 and 2.13). Proving the transference result then consists essentially in showing that X_2 models every filtered nilmanifold of degree at most 2 (see Theorem 2.15 and Proposition 2.16). To prove the latter statement, we use in particular a certain decomposition of general filtered nilmanifolds of degree at most 2, which we establish in Section 3 (see Proposition 3.2).

In Section 6, we rule out an a-priori plausible approach to generalizing Theorem 1.5 for higher-complexity systems, by showing that the degree-3 filtered torus $(\mathbb{T}^3, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)} \times \mathbb{R}_{\bullet(3)})$ does not model all filtered nilmanifolds of degree at most 3. The counterexample involves the Heisenberg nilmanifold with a certain degree-3 filtration (see Proposition 6.1). This example shows that the possibility to use a torus of this kind in a transference result such as Theorem 2.1 is a specific feature of systems of complexity at most 2. This phenomenon provides a new illustration of the marked increase in difficulty in passing from systems of complexity 2 to systems of complexity at least 3 (this increase has been observed in different contexts before, for example in [8, §11]).

2. A TRANSFERENCE RESULT FOR SYSTEMS OF COMPLEXITY 2

To prove the convergence of quantities such as $m_k(\alpha, \mathbb{Z}_p)$, it is natural to try to transfer a function on a large group \mathbb{Z}_p to another such group \mathbb{Z}_q while keeping certain Ψ -averages of this function roughly unchanged. This idea is already present in Croot's proof of Theorem 1.1 in [3], where it is implemented by an argument that uses Fourier analysis and that is extendable to any system of complexity 1 (in [3] it is used just for the central case of 3-term arithmetic progressions).

Extending this method towards the continuous setting, results such as Theorem 1.2 can be deduced from transference results that are similar except that now one has to allow either of \mathbb{Z}_p or \mathbb{Z}_q to be replaced with \mathbb{T} . This was originally done in [20], also using Fourier analysis (see [20, Proposition 4.2.8], and also [1, Corollary 3.8]).

Theorem 1.5 can be obtained in a similar spirit as an immediate consequence of the following result, but our proof of this result will use *quadratic* Fourier analysis.

Theorem 2.1 (Transference for systems of complexity 2 between \mathbb{Z}_p and X_2).

For every $\epsilon > 0$, there exists $C > 0$ such that the following holds. Let $X = \mathbb{Z}_p$ with $p \geq C$ and let $X' = X_2$. Then for every function $f : X \rightarrow [0, 1]$, there exists a continuous function $f' : X' \rightarrow [0, 1]$ such that, for every system Ψ of linear forms of size⁶ at most $1/\epsilon$ and complexity at most 2, we have $|S_\Psi(f : X) - S_\Psi(f' : X')| \leq \epsilon$. The last sentence holds also with $X = X_2$ (with f being measurable) and $X' = \mathbb{Z}_p$.

We split the proof into two cases. The main case $X = \mathbb{Z}_p$, $X' = X_2$ will occupy most of the sequel. The other case, namely $X = X_2$, $X' = \mathbb{Z}_p$, is simpler and is treated briefly in Section 4; see Proposition 4.3. We shall end this section with an outline of the proof of the main case, in Subsection 2.2. Before that, however, we pause to give a sense in which the nilmanifold X_2 in Theorem 2.1 cannot be simplified.

2.1. Some counterexamples.

One may call a filtered nilmanifold playing the role of X_2 in Theorem 2.1 a *continuous model* for systems of complexity at most 2 on groups \mathbb{Z}_p . One might then wonder whether simpler models than X_2 could be used, and in particular whether one of the two components of X_2 , namely $(\mathbb{T}, \mathbb{R}_{\bullet(1)})$, $(\mathbb{T}, \mathbb{R}_{\bullet(2)})$, could already suffice. In this subsection we show that, if in Theorem 2.1 we allow complex-valued functions and also averages involving complex conjugation, then the theorem does not hold with any of these two components.

We start by ruling out $(\mathbb{T}, \mathbb{R}_{\bullet(1)}) = X_1$. To this end we shall use the systems of linear forms related to the Gowers norms $\|\cdot\|_{U^2}$ and $\|\cdot\|_{U^3}$. Let

$$\Psi_{U^2} : n \in \mathbb{Z}^3 \mapsto (n_1, n_1 + n_2, n_1 + n_3, n_1 + n_2 + n_3), \quad (5)$$

and let

$$\begin{aligned} \Psi_{U^3} : n \in \mathbb{Z}^4 \mapsto & (n_1, n_1 + n_2, n_1 + n_3, n_1 + n_2 + n_3, \\ & n_1 + n_4, n_1 + n_2 + n_4, n_1 + n_3 + n_4, n_1 + n_2 + n_3 + n_4). \end{aligned} \quad (6)$$

For a bounded measurable function $f : Z \rightarrow \mathbb{C}$ on a compact abelian group Z , we set

$$S'_{\Psi_{U^2}}(f : Z) = \int_{Z^3} f(x_1) \overline{f(x_1 + x_2)} \overline{f(x_1 + x_3)} f(x_1 + x_2 + x_3) \, d\mu_{Z^3}(x_1, x_2, x_3),$$

and we define $S'_{\Psi_{U^3}}(f : Z)$ similarly (in particular $S'_{\Psi_{U^d}}(f : Z) = \|f\|_{U^d(Z)}^2$, $d = 2, 3$).

We then have the following result, which shows that Theorem 2.1 does not hold for all functions taking values in the unit disc $\mathcal{D} \subseteq \mathbb{C}$ if we replace X_2 with X_1 .

⁶We recall this notion of size in Definition A.7.

Proposition 2.2. *For each prime p , let $f_p : \mathbb{Z}_p \rightarrow \mathcal{D}$, $x \mapsto e(x^2/p) = \exp(2\pi i x^2/p)$. There exists an absolute constant $\delta > 0$ such that, for every p sufficiently large, there is no continuous function $f : \mathbb{T} \rightarrow \mathcal{D}$ such that*

$$\max \left\{ \left| S'_{\Psi_{U^2}}(f_p : \mathbb{Z}_p) - S'_{\Psi_{U^2}}(f : \mathbb{T}) \right|, \left| S'_{\Psi_{U^3}}(f_p : \mathbb{Z}_p) - S'_{\Psi_{U^3}}(f : \mathbb{T}) \right| \right\} < \delta. \quad (7)$$

The function f_p in this result has been used before as a source of examples in the discrete setting (see [8, §4]). In particular, a standard calculation shows that $S'_{\Psi_{U^3}}(f_p : \mathbb{Z}_p) = 1$ while $S'_{\Psi_{U^2}}(f_p : \mathbb{Z}_p) = o(1)_{p \rightarrow \infty}$ (see [23, Exercise 11.1.12]). We want to show that this behaviour cannot be reproduced by a continuous \mathcal{D} -valued function on \mathbb{T} . This will follow from a combination of two results.

Given compact abelian groups Z, Z' , we say that a continuous map $P : Z \rightarrow Z'$ is a *polynomial map of degree $\leq d$* if

$$\Delta_{h_1} \cdots \Delta_{h_{d+1}} P(x) = 0 \text{ for every } x, h_1, \dots, h_{d+1} \in Z, \quad (8)$$

where $\Delta_h P(x) := P(x + h) - P(x)$.

There is a related definition of polynomial maps between filtered groups, that we recall in Appendix A; see Definition A.2, after which we also recall the relation between the two definitions.

Of the two results that we shall combine as mentioned above, the first is the following characterization, due to Eisner and Tao, of bounded functions that are nearly extremal for the U^d norm [7, Theorem 1.1].

Theorem 2.3 (L^∞ near-extremisers on compact abelian groups). *Let d be a positive integer, let Z be a compact abelian group, and let $f \in L^\infty(Z)$ be such that $\|f\|_{L^\infty(Z)} \leq 1$. If $\|f\|_{U^d(Z)} \geq 1 - \epsilon$, then there exists a continuous polynomial map $P : Z \rightarrow \mathbb{R}/\mathbb{Z}$ of degree $\leq d - 1$ such that $\|f - e(P)\|_{L^1(Z)} = o(1)_{\epsilon \rightarrow 0}$.*

The second result concerns continuous polynomial maps between tori. (A special case appeared already in work of the second named author; see [22, Lemma 1.3].)

Lemma 2.4. *Let $m \geq 1$ and fix any filtrations $\mathbb{T}_\bullet^m, \mathbb{T}_\bullet$ on \mathbb{T}^m, \mathbb{T} . Then every continuous polynomial map $f : (\mathbb{T}^m, \mathbb{T}_\bullet^m) \rightarrow (\mathbb{T}, \mathbb{T}_\bullet)$ is a polynomial map $\mathbb{T}^m \rightarrow \mathbb{T}$ of degree ≤ 1 .*

Proof. Note first that the map f , which is assumed to be polynomial in the sense of Definition A.2, is then also polynomial of degree at most d in the sense of (8), where d is the degree of the filtration \mathbb{T}_\bullet . Now suppose for a contradiction that f is not of degree ≤ 1 . Then by repeated applications of difference operators Δ_h to f , we eventually obtain a map g such that $\Delta_{h_3} \Delta_{h_2} \Delta_{h_1} g(x) = 0$ for all $x, h_1, h_2, h_3 \in \mathbb{T}^m$, but $\Delta_{h_2} \Delta_{h_1} g(x)$ is not

zero for all x, h_1, h_2 . It therefore suffices to obtain a contradiction if f is such a map. In other words, we may suppose that f is polynomial of degree ≤ 2 but not of degree ≤ 1 . Now note that, in this case, for each fixed h the map $\Delta_h f$ is polynomial of degree ≤ 1 . Setting $\theta_h = \Delta_h f(0)$, we therefore have that $x \mapsto \Delta_h f(x) - \theta_h$ is a continuous homomorphism $\mathbb{T}^m \rightarrow \mathbb{T}$. Hence there is $n_h \in \mathbb{Z}^m$ such that $\Delta_h f(x) = n_h \cdot x + \theta_h$. (Here $n \cdot x = n(1)x(1) + \dots + n(m)x(m)$.) On the other hand, the uniform continuity of f implies that the map $h \mapsto \Delta_h f - \theta_h$ is continuous from \mathbb{T}^m to the space of continuous functions $\mathbb{T}^m \rightarrow \mathbb{T}$. (On the latter space we use the topology given by the metric $d(f, g) = \sup_{x \in \mathbb{T}^m} \|f(x) - g(x)\|_{\mathbb{T}}$.) The continuous homomorphisms $\mathbb{T}^m \rightarrow \mathbb{T}$, being of the form $x \mapsto n \cdot x$ for some $n \in \mathbb{Z}^m$, form a discrete set in this space. Since \mathbb{T}^m is connected, the map $h \mapsto n_h$ must therefore be constant, so there is $n \in \mathbb{Z}^m$ such that $\Delta_h f(x) = n \cdot x + \theta_h$ for all h, x . Applying this with $h = 0$ implies that $n \cdot x + \theta_0 = 0$ for all $x \in \mathbb{T}^m$, which implies that $n = 0$. Thus $\Delta_h f(x) = \theta_h$ for all h, x , and so $\Delta_k \Delta_h f(x) = 0$ for all h, k, x , which contradicts that f is not polynomial of degree ≤ 1 . \square

Proof of Proposition 2.2. If $f : \mathbb{T} \rightarrow \mathcal{D}$ is a continuous function satisfying

$$\left| S'_{\Psi_{U^3}}(f_p : \mathbb{Z}_p) - S'_{\Psi_{U^3}}(f : \mathbb{T}) \right| < \delta,$$

then by Theorem 2.3 together with the fact that $S'_{\Psi_{U^3}}(f_p : \mathbb{Z}_p) = 1$, there is some continuous polynomial map $P : \mathbb{T} \rightarrow \mathbb{T}$ of degree ≤ 2 such that $\|f - e(P)\|_{L^1(\mathbb{T})} \leq \epsilon$, where $\epsilon = \epsilon(\delta) \rightarrow 0$ as $\delta \rightarrow 0$. By Lemma 2.4 we have in fact that P is of degree ≤ 1 . But then a simple calculation shows that $S'_{\Psi_{U^2}}(f : \mathbb{T}) = S'_{\Psi_{U^2}}(e(P) : \mathbb{T}) - o(1) = 1 - o(1)$ as $\delta \rightarrow 0$. Since $S'_{\Psi_{U^2}}(f_p : \mathbb{Z}_p) = o(1)_{p \rightarrow \infty}$, the inequality $\left| S'_{\Psi_{U^2}}(f_p : \mathbb{Z}_p) - S'_{\Psi_{U^2}}(f : \mathbb{T}) \right| < \delta$ fails for δ sufficiently small and p sufficiently large. \square

Remark 2.5. Ruling out $Y = (\mathbb{T}, \mathbb{R}_{\bullet(2)})$ is a much simpler task. It suffices to use the trivial form $\Psi_0 : n \in \mathbb{Z} \mapsto n$ (note that $S_{\Psi_0}(f : Z) = \int_Z f d\mu_Z$ for any compact abelian group Z) and some system of complexity 1, say $\Psi : n \in \mathbb{Z}^2 \mapsto (n_1, n_2, n_1 + n_2)$. Indeed, we can find a set $A_p \subseteq \mathbb{Z}_p$ such that for large p the following inequality fails for every continuous function $f : \mathbb{T} \rightarrow [0, 1]$:

$$\max \left\{ \left| S_{\Psi_0}(1_{A_p} : \mathbb{Z}_p) - S_{\Psi_0}(f : Y) \right|, \left| S_{\Psi}(1_{A_p} : \mathbb{Z}_p) - S_{\Psi}(f : Y) \right| \right\} < 1/100.$$

To see this, note first that $Y^\Psi = \mathbb{T}^3$ (using Definitions A.1 and A.6), so any such function f satisfies $S_\Psi(f : Y) = (\int_{\mathbb{T}} f)^3 = S_{\Psi_0}(f : Y)^3$. By contrast, for a function $f_p : \mathbb{Z}_p \rightarrow [0, 1]$, the average $S_\Psi(f_p : \mathbb{Z}_p) = \mathbb{E}_{n_1, n_2 \in \mathbb{Z}_p} f(n_1)f(n_2)f(n_1 + n_2)$ can be much smaller than $S_{\Psi_0}(f_p : \mathbb{Z}_p)^3$. For instance, let $f_p = 1_{A_p}$ where A_p is the large sumfree set $(p/3, 2p/3) \subseteq \mathbb{Z}_p$.

Remark 2.6. It would be interesting to establish stronger variants of Proposition 2.2 that would rule out X_1 as a continuous model for systems of complexity 2 more decisively. For instance, one can ask whether in that proposition $e(x^2/p)$ can be replaced with some $[0, 1]$ -valued function. Such a real-valued counterexample could require allowing more than two systems of complexity at most 2 in the maximum in (7). In closer relation to Theorem 1.5, it would be conclusive to find a system Ψ of complexity 2 for which $\lim_{p \rightarrow \infty} m_\Psi(\alpha, \mathbb{Z}_p) < m_\Psi(\alpha, X_1)$ for some α . In particular, we do not know whether this inequality holds for the system Ψ_4 of 4-term arithmetic progressions.

Question 2.7. *Does the inequality $m_{\Psi_4}(\alpha, X_2) < m_{\Psi_4}(\alpha, X_1)$ hold for some $\alpha \in [0, 1]$?*

As a counterpoint to the last remark, let us mention that in the next subsection we shall introduce a definition of modelling, concerning filtered nilmanifolds, and that with respect to this definition X_2 can definitely not be replaced with X_1 (as detailed in Remark 2.17 below). This definition is in some sense more natural than (though related to) the notion of a continuous model for systems on \mathbb{Z}_p considered above.

2.2. A reduction of the main case of Theorem 2.1.

Our central aim from now on is to establish the following result.

Proposition 2.8. *For every $\epsilon > 0$, there exists $C > 0$ such that the following holds. For every function $f : \mathbb{Z}_p \rightarrow [0, 1]$ with $p \geq C$, there exists a continuous function $f' : \mathbb{T} \times \mathbb{T} \rightarrow [0, 1]$ such that, for every system Ψ of linear forms of size at most $1/\epsilon$ and complexity at most 2, we have $|S_\Psi(f' : X_2) - S_\Psi(f : \mathbb{Z}_p)| \leq \epsilon$.*

Remark 2.9. This result is equivalent to the seemingly weaker version in which f' is only claimed to be measurable. (The version in terms of measurable functions is also natural in that it accommodates the indicator functions of measurable sets.) The equivalence can be seen using Lusin's theorem to approximate f' in $L^1(\mathbb{T}^2)$ by a continuous function h , and then the fact (which will also be used later) that for all measurable functions f', h on \mathbb{T}^2 bounded by 1 and every system $\Psi : \mathbb{Z}^D \rightarrow \mathbb{Z}^t$ of finite complexity, we have

$$|S_\Psi(f' : X_2) - S_\Psi(h : X_2)| \leq t \|f' - h\|_{L^1(\mathbb{T}^2)}.$$

This L^1 -continuity of $S_\Psi(\cdot : X_2)$ follows from the fact that any coordinate projection $(\mathbb{T}^2)^t \rightarrow \mathbb{T}^2$ restricts to a surjective map from X_2^Ψ to \mathbb{T}^2 . (Note that the complexity assumption for Ψ implies that none of the linear forms is the 0-form; see Definition A.7 and the paragraph thereafter.)

For a continuous function $F : X \rightarrow \mathbb{C}$ on a metric space (X, d) , we define the Lipschitz norm of F by $\|F\|_{\text{Lip}} = \|F\|_{\infty} + \sup_{x, y \in X, x \neq y} \frac{|F(x) - F(y)|}{d(x, y)}$.

Definition 2.10 (Quantitative equidistribution). Let (X, μ) be a probability space and let X' be a metric space with a Borel probability measure μ' . We say that a measurable map $\phi : X \rightarrow X'$ is δ -*equidistributed* if for every Lipschitz function $F : X' \rightarrow \mathbb{C}$ we have

$$\left| \int_X F \circ \phi \, d\mu - \int_{X'} F \, d\mu' \right| \leq \delta \|F\|_{\text{Lip}}. \quad (9)$$

On a filtered nilmanifold with a fixed Mal'cev basis \mathcal{X} (see Definition A.4) there is a convenient metric, which was defined in [12, Definition 2.2]. Using this metric we shall now define a notion of equidistribution concerning certain maps between filtered nilmanifolds, which plays a central role in the sequel. This is a strong form of equidistribution in the sense that, in addition to the given map being equidistributed, we require various multiparameter versions of the map, corresponding to Leibman nilmanifolds for systems of linear forms, also to be equidistributed. In the special case of polynomial sequences, this is related to the stronger notion of *irrationality* from [13] (see also [2]).

Given filtered groups $(G, G_{\bullet}), (G', G'_{\bullet})$, a polynomial map $g \in \text{poly}_0(G_{\bullet}, G'_{\bullet})$ (see Definition A.2), and subgroups Γ, Γ' of G, G' respectively, we say that g is (Γ, Γ') -*consistent* if for every $\gamma \in \Gamma$ the map $x \mapsto g(x)^{-1}g(x\gamma)$ is Γ' -valued on G . Note that if g is (Γ, Γ') -consistent then $x\Gamma \mapsto g(x)\Gamma'$ is a well-defined map $G/\Gamma \rightarrow G'/\Gamma'$. We denote this map by ϕ_g and say that it is *induced* by g . Note also the following composition property: if $g \in \text{poly}_0(G_{\bullet}, G'_{\bullet})$ is (Γ, Γ') -consistent and $g' \in \text{poly}_0(G'_{\bullet}, G''_{\bullet})$ is (Γ', Γ'') -consistent, then $g' \circ g$ is in $\text{poly}_0(G_{\bullet}, G''_{\bullet})$ (by composition of polynomial maps; see the paragraph after Definition A.2) and is (Γ, Γ'') -consistent.

The strong notion of equidistribution is the following.

Definition 2.11 (Balanced map). Let $X = (G/\Gamma, G_{\bullet})$ be a possibly disconnected filtered nilmanifold (see Remark A.5), let $X' = (G'/\Gamma', G'_{\bullet}, \mathcal{X}')$ be a filtered based nilmanifold (see Definition A.4), and let $g : G \rightarrow G'$ be a continuous (Γ, Γ') -consistent map in $\text{poly}_0(G_{\bullet}, G'_{\bullet})$. For $\delta > 0$ we say that $\phi_g : X \rightarrow X'$ is δ -*balanced* if for each system $\Psi : \mathbb{Z}^D \rightarrow \mathbb{Z}^t$ of linear forms of size at most $1/\delta$ the following map is δ -equidistributed:⁷

$$\phi_g^t : X^{\Psi} \rightarrow X'^{\Psi}, \quad x\Gamma^t \mapsto g^t(x)\Gamma'^t = (g(x_1), \dots, g(x_t))\Gamma'^t. \quad (10)$$

⁷There is a slight abuse of notation in (10) in that we are implicitly using the identification of a nilmanifold X^{Ψ} with a subnilmanifold of X^t , via the embedding $x\Gamma^{\Psi} \mapsto x\Gamma^t$. Note also that the metric used on X'^{Ψ} for δ -equidistribution here is the restriction of the metric on X'^t related to the Mal'cev basis \mathcal{X}' ; see [2, Appendix A].

Remark 2.12. We allow X to be possibly disconnected here (in the sense of Remark A.5) so that the notion of a balanced map can concern maps defined on a filtered finite abelian group X . Thus, for instance, if $G = \mathbb{Z}$ with the lower central series and $\Gamma = p\mathbb{Z}$, we have that $g : G \rightarrow G'$ is (Γ, Γ') -consistent if and only if g is a polynomial sequence in $\text{poly}_0(\mathbb{Z}, G'_\bullet)$ that is *p-periodic mod Γ'* in the sense of [2, Definition 2.10]. In this case, note that ϕ_g is δ -balanced if the polynomial g is sufficiently *irrational* in the sense of [2, Definition 4.7].

This notion of a balanced map is related to that of a *balanced morphism*, a natural analogue in the category of nilspaces (see [22, Definition 1.2]); we shall not detail this relation in this paper.

In Definition 2.11 we are implicitly using the property that $g^t(x)$ is in the Leibman group $(G', G'_\bullet)^\Psi$ whenever $x \in (G, G_\bullet)^\Psi$. This property is nontrivial but a short proof can be given using known facts (see Proposition A.3).

Note also that if $\phi_g : X \rightarrow X'$ is δ -balanced then in particular, using the relation between the metrics corresponding to \mathcal{X}' and \mathcal{X}'^t (see [2, Lemmas A.3 and A.4]) we have, for every Lipschitz function $F : X' \rightarrow \mathbb{C}$,

$$|S_\Psi(F \circ \phi_g : X) - S_\Psi(F : X')| \leq \delta \|F^{\otimes t}\|_{\text{Lip}(\mathcal{X}'^t)} \leq \delta t \|F\|_{\text{Lip}(\mathcal{X}')}. \quad (11)$$

(Here $F^{\otimes t}$ denotes the function $X^t \rightarrow \mathbb{C}$, $(x_1, \dots, x_t)\Gamma^t \mapsto F(x_1\Gamma) \cdots F(x_t\Gamma)$.)

With the notion of balanced maps we can now give one of the central definitions of this paper.

Definition 2.13 (Modelling filtered nilmanifolds). Let X, X' be filtered nilmanifolds. We say that X *models* X' if there exists a Mal'cev basis \mathcal{X}' on X' such that for every $\delta > 0$ there exists a δ -balanced map $\phi_g : X \rightarrow X'$.

The basis \mathcal{X}' is used just to have the metric structure underlying the notion of balanced map, but note that if the main statement holds for one such basis then it holds for any other. One can see this using results relating the metrics corresponding to two given bases; see for instance [12, Lemma A.17]. Note also that modelling induces a preorder on the set of filtered nilmanifolds. We discuss this further at the end of Section 7.

The first main tool that we use to prove Proposition 2.8 is the following consequence of the regularity method for the Gowers U^3 norm.

Theorem 2.14. *For every $\delta > 0$ there exists $C > 0$ such that the following holds. For every function $f : \mathbb{Z}_p \rightarrow [0, 1]$ with $p \geq C$, there is a filtered based nilmanifold $X = (G/\Gamma, G_\bullet, \mathcal{X})$ of degree at most 2 and complexity at most C , and a continuous*

function $F : G/\Gamma \rightarrow [0, 1]$ with $\|F\|_{\text{Lip}(\mathcal{X})} \leq C$, such that for every system Ψ of integer linear forms of size at most $1/\delta$ and complexity at most 2, we have

$$|S_\Psi(f : \mathbb{Z}_p) - S_\Psi(F : X)| \leq \delta. \quad (12)$$

This was essentially proved in [2] using arguments and results from [13, 22], but for the above version we need some small modifications, so for completeness we include a proof in Appendix A (see Theorem A.8).

Given Theorem 2.14, proving Proposition 2.8 reduces to establishing the following.

Theorem 2.15. *Let $X_2 = (\mathbb{T}^2, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)})$ and let X be a filtered nilmanifold of degree at most 2. Then X_2 models X .*

Proposition 2.16. *Theorem 2.15 implies Proposition 2.8.*

Proof. Given $\epsilon > 0$ as in Proposition 2.8, we first apply Theorem 2.14, with parameter δ_1 a function of ϵ to be fixed later. Thus for every $f : \mathbb{Z}_p \rightarrow [0, 1]$ with $p \geq C(\delta_1)$ there exists $F : X \rightarrow [0, 1]$ with $\|F\|_{\text{Lip}(\mathcal{X})} \leq C(\delta_1)$ such that for every system $\Psi : \mathbb{Z}^D \rightarrow \mathbb{Z}^t$ of complexity at most 2 and size at most $1/\delta_1$, we have $|S_\Psi(f : \mathbb{Z}_p) - S_\Psi(F : X)| \leq \delta_1$.

Now, by Theorem 2.15, there exists a δ_2 -balanced map $\phi_g : X_2 \rightarrow X$ (with respect to the metric on X given by \mathcal{X}), where $\delta_2 > 0$ is a function of ϵ that we fix later.

Let $f' : X_2 \rightarrow [0, 1]$ be the continuous function $F \circ \phi_g$. Then by (11) we have

$$|S_\Psi(F : X) - S_\Psi(f' : X_2)| \leq \delta_2 t \|F\|_{\text{Lip}(\mathcal{X})}^t \leq \delta_2 \delta_1^{-1} C(\delta_1)^{1/\delta_1}.$$

Letting $\delta_1 = \epsilon/2$ and δ_2 such that $\delta_2 \delta_1^{-1} C(\delta_1)^{1/\delta_1} \leq \epsilon/2$, the result follows. \square

To prove Theorem 2.15, we shall first replace X with a product of two simpler filtered nilmanifolds, namely a 2-step nilmanifold Y_1 with lower central series and a torus Y_2 with the maximal degree-2 filtration. This is done in the next section. As a consequence, finding the desired δ -balanced map ϕ_g for Theorem 2.15 will be reduced to finding a sufficiently balanced map $\phi_1 : X_1 \rightarrow Y_1$ and another such map $\phi_2 : (\mathbb{T}, \mathbb{R}_{\bullet(2)}) \rightarrow Y_2$ (see Proposition 3.6). We obtain these maps in Sections 4 and 5 respectively.

Remark 2.17. As mentioned at the end of Subsection 2.1, in Theorem 2.15 one cannot replace X_2 with X_1 . In fact X_1 does not even model $Y = (\mathbb{T}, \mathbb{R}_{\bullet(2)})$. To see this, let $\Psi = \Psi_{U^2}$ as in (5). Computing the Leibman nilmanifolds, we find that $Y^\Psi = \mathbb{T}^4$, whereas X_1^Ψ is the 3-dimensional subtorus $\Psi(\mathbb{T}^3)$ of \mathbb{T}^4 . Since a continuous polynomial map $\phi_g : \mathbb{T} \rightarrow \mathbb{T}$ with $\phi_g(0) = 0$ must be a homomorphism (by Lemma 2.4), we have $\phi_g^4(\Psi(\mathbb{T}^3)) \subseteq \Psi(\mathbb{T}^3)$. Therefore, for δ sufficiently small, for every such map ϕ_g we have that ϕ_g^4 is not δ -equidistributed in Y^Ψ , and so $\phi_g : X_1 \rightarrow Y$ is not δ -balanced.

3. A PRODUCT DECOMPOSITION OF DEGREE-2 FILTERED NILMANIFOLDS

The main result of this section describes the structure of a general filtered nilmanifold of degree 2 by decomposing it into the product of two nilmanifolds with very specific filtrations. The formal statement uses the following notions.

Definition 3.1 (Isomorphism of filtered nilmanifolds). Let $X = (G/\Gamma, G_\bullet)$ and $X' = (G'/\Gamma', G'_\bullet)$ be filtered nilmanifolds of degree d . We say that X and X' are *isomorphic* if there is a map $\theta : G \rightarrow G'$ satisfying the following conditions:

- (i) θ is an isomorphism of filtered Lie groups. This means that for each $i \geq 0$ the restriction of θ to $G_{(i)}$ is an isomorphism of Lie groups $G_{(i)} \rightarrow G'_{(i)}$.
- (ii) We have $\theta(\Gamma) = \Gamma'$.

If Mal'cev bases $\mathcal{X} = \{x_1, \dots, x_r\}$, $\mathcal{X}' = \{x'_1, \dots, x'_s\}$ are given on X, X' respectively, then we say that X, X' are *isomorphic* as filtered *based* nilmanifolds if there is a map $\theta : G \rightarrow G'$ satisfying condition (i) above (so that in particular we have $r = s$) and satisfying also the following condition:

- (ii') For each $j \in [r]$ we have $\theta(\exp(x_j)) = \exp(x'_j)$.

Note that (i) and (ii') together imply (ii), since by the defining properties of Mal'cev bases we have $\Gamma = \{\exp(t_1 x_1) \cdots \exp(t_m x_m) : t_j \in \mathbb{Z}\}$.

The structural result can be stated as follows.

Proposition 3.2. *Let $X = (G/\Gamma, G_\bullet)$ be a filtered nilmanifold of degree at most 2. Then X is isomorphic to $Y_1 \times Y_2$, where Y_1 is a 2-step nilmanifold with lower central series, and Y_2 is a torus with the maximal degree-2 filtration.*

We shall actually prove a refinement of this result, namely Proposition 3.3 below, in which an additional metric structure on X given by a fixed Mal'cev basis is also approximately conserved by the isomorphism. This refinement can be useful from a quantitative point of view, especially in relation to the regularity lemma for the U^3 norm (see Remark 3.4).

Given two Mal'cev bases $\mathcal{X}, \mathcal{X}'$ on a filtered nilmanifold, we say that \mathcal{X}' is *Q-rational relative to \mathcal{X}* if each element from \mathcal{X}' is a linear combination of the elements of \mathcal{X} with rational coefficients of height at most Q . (Similar terminology was used in [12, Definitions 2.4, 2.5].)

Proposition 3.3. *There is a function $\eta : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ such that the following holds. Let $Q \geq 2$, and let $X = (G/\Gamma, G_\bullet, \mathcal{X})$ be of degree 2 and complexity at most Q . Then there exists a Mal'cev basis \mathcal{X}' for $(G/\Gamma, G_\bullet)$ that is $\eta(Q)$ -rational relative to \mathcal{X} , such that $X' := (G/\Gamma, G_\bullet, \mathcal{X}')$ is isomorphic (as a filtered based nilmanifold) to $Y_1 \times Y_2$, where $Y_1 = (H/\Gamma_H, H_{\bullet(1)}, \mathcal{X}_H)$ is a 2-step nilmanifold with lower-central series and complexity at most $\eta(Q)$, and $Y_2 = (\mathbb{T}^m, \mathbb{R}_{\bullet(2)}^m)$ with the standard basis on \mathbb{R}^m and with $m \leq Q$.*

As the proof reveals, we can take $\eta(Q) = Q^K$ for some constant K depending only on the dimension and degree of (G, G_\bullet) ; in particular, K also depends only on Q .

Proof. In this proof, just as in [12, Appendix A], the constants implicit in $O(1)$ notations are allowed to depend on the dimension and degree of (G, G_\bullet) .

The basis \mathcal{X} is adapted to the given filtration $G_\bullet = (G_{(0)}, G_{(1)}, G_{(2)}, \{\text{id}_G\}, \dots)$. First, we claim that there is a Mal'cev basis \mathcal{X}' for G/Γ , which is $Q^{O(1)}$ -rational relative to \mathcal{X} , and which refines \mathcal{X} in the sense that \mathcal{X}' is also adapted to G_\bullet and in addition it passes through the subgroup $G_{(3)} := [G, G] \leq G_{(2)}$.

To see this, note first that $G_{(3)}$ is a Q -rational subgroup of G relative to \mathcal{X} , in the sense of [12, Definition 2.5]. Indeed, letting \mathfrak{g} denote the Lie algebra of G , we have the fact that the Lie subalgebra $\mathfrak{g}_{(3)}$ corresponding to $G_{(3)}$ is spanned by Lie brackets $[x_i, x_j]$ where $x_i, x_j \in \mathcal{X}$. (By [5, Proposition 5.2.1] we know that $\mathfrak{g}_{(3)}$ is the \mathbb{R} -span of elements $[v, w]$, $v, w \in \mathfrak{g}$. Since v, w are themselves linear combinations of the x_i , the fact follows.) By definition of \mathcal{X} being Q -rational (see [12, Definition 2.4]), each $[x_i, x_j]$ is a Q -rational combination of the x_i , so these rational combinations span $\mathfrak{g}_{(3)}$, and so we can find a basis of such combinations for $\mathfrak{g}_{(3)}$, whence $G_{(3)}$ is indeed Q -rational relative to \mathcal{X} . Next, note that since $G'_\bullet := (G_{(0)}, G_{(1)}, G_{(2)}, G_{(3)}, \{\text{id}_G\}, \dots)$ is a filtration on G , we may apply [12, Proposition A.10] with $G' = G$ and the filtrations G_\bullet, G'_\bullet , and thus obtain the claimed basis \mathcal{X}' .

Now we use \mathcal{X}' to define a 2-step nilpotent Lie subgroup H of G and an abelian Lie subgroup V , which will yield our decomposition. Letting m_i denote the dimension of $G_{(i)}$, for $i \in [3]$ (and $m_0 = \dim(G)$), we define first

$$\begin{aligned} H &= e^{\mathbb{R} \cdot x_1} \dots e^{\mathbb{R} \cdot x_{m_0-m_2}} \cdot e^{\mathbb{R} \cdot x_{m_0-m_3+1}} \dots e^{\mathbb{R} \cdot x_{m_0}} \\ &= e^{\mathbb{R} \cdot x_1} \dots e^{\mathbb{R} \cdot x_{m_0-m_2}} \cdot [G, G]. \end{aligned}$$

Thus, letting π denote the quotient map $G \rightarrow G/[G, G]$, we have that H is the preimage under π of the \mathbb{R} -span of the vectors $\pi(e^{x_1}), \dots, \pi(e^{x_{m_0-m_2}})$. In particular H is a normal Lie subgroup of G (see for instance [19, Theorem 3.5, p. 18]), connected and simply-connected.

Next, we define

$$V = e^{\mathbb{R} \cdot x_{m_0-m_2+1}} \dots e^{\mathbb{R} \cdot x_{m_0-m_3}}.$$

This is an abelian Lie group (connected and simply-connected), indeed it is just the vector subspace of $G_{(2)}$ spanned by the elements e^{x_j} , $j \in (m_0 - m_2, m_0 - m_3]$.

These normal Lie subgroups H, V of G satisfy $H \cap V = \{\text{id}\}$ and $H \cdot V = G$. In fact, we have an isomorphism of Lie groups $\theta : G \rightarrow H \times V$ defined by $\theta(g) = (h, v)$ where $g = e^{r_1 x_1} \dots e^{r_{m_0} x_{m_0}}$, $h = e^{r_1 x_1} \dots e^{r_{m_0-m_2} x_{m_0-m_2}} e^{r_{m_0-m_3+1} x_{m_0-m_3+1}} \dots e^{r_{m_0} x_{m_0}}$, and $v = e^{r_{m_0-m_2+1} x_{m_0-m_2+1}} \dots e^{r_{m_0-m_3} x_{m_0-m_3}}$.

We claim that θ is an isomorphism of filtered Lie groups if we endow $H \times V$ with the filtration $H_{\bullet(1)} \times V_{\bullet(2)}$. To prove this we just need to show that $\theta(G_{(2)}) = [H, H] \times V$. Since $G_{(2)} = e^{\mathbb{R} \cdot x_{m_0-m_2+1}} \dots e^{\mathbb{R} \cdot x_{m_0}}$, by definition of V we have that $\theta(G_{(2)}) = [G, G] \times V$. We also have $[G, G] = [H, H]$. Indeed, on one hand we have $G \subseteq H \cdot G_{(2)}$, and on the other hand since G is 2-step nilpotent we have that $G_{(2)}$ lies in the center of G and the commutator map satisfies $[g_1 g_2, g_3] = [g_1, g_3] \cdot [g_2, g_3]$, whence $[G, G] \subseteq [H \cdot G_{(2)}, H \cdot G_{(2)}] \subseteq [H, H]$. This proves our claim.

Note that by construction H and V are rational subgroups of G and so by [5, Theorem 5.1.11] we have that the subgroups $\Gamma_H := \Gamma \cap H$ and $\Gamma_V := \Gamma \cap V$ are lattices in H, V respectively. Moreover, the following are then Mal'cev bases for $(H/\Gamma_H, H_{\bullet(1)})$ and $(V/\Gamma_V, V_{\bullet(2)})$ respectively:

$$\mathcal{X}_H = \{x_1, \dots, x_{m_0-m_2}, x_{m_0-m_3+1}, \dots, x_{m_0}\}, \quad \mathcal{X}_V = \{x_{m_0-m_2+1}, \dots, x_{m_0-m_3}\},$$

where \mathcal{X}_H is $Q^{O(1)}$ rational (and \mathcal{X}_V is 0-rational). Letting $\mathfrak{h}, \mathfrak{v}$ denote the Lie algebras of H, V , the Lie algebra of $H \times V$ is $\mathfrak{h} \oplus \mathfrak{v}$, and then the Mal'cev basis $\mathcal{X}_{H \times V}$ for

$$((H \times V)/(\Gamma_H \times \Gamma_V), H_{\bullet(1)} \times V_{\bullet(2)})$$

satisfying condition (ii') in Definition 3.1 is

$$\{(x_1, 0), \dots, (x_{m_0-m_2}, 0), (0, x_{m_0-m_2+1}), \dots, (0, x_{m_0-m_3}), (x_{m_0-m_3+1}, 0), \dots, (x_{m_0}, 0)\}.$$

Since $(V/\Gamma_V, V_{\bullet(2)}, \mathcal{X}_V)$ is isomorphic to $(\mathbb{T}^m, \mathbb{R}_{\bullet(2)}^m)$ with the standard basis, the proof is complete. \square

Throughout the sequel, the Mal'cev basis on a torus \mathbb{T}^m is by default the standard basis on \mathbb{R}^m , so we shall not specify the basis on filtered tori from now on.

Remark 3.4. The regularity lemma for the $U^3(\mathbb{Z}_p)$ norm (stated for instance in [2, Theorem 5.1]) can be refined using Proposition 3.3, by replacing the unspecified nilmanifold X of degree 2 in that result by a product nilmanifold of the form $Y_1 \times Y_2$

given by the proposition. Note that in doing so it is important to control the complexity (or quantitative rationality) of the basis on $Y_1 \times Y_2$ only in terms of the complexity of X , so that the metric structures, governing Lipschitz constants etc., can be related in a way that depends only on the complexity of X . The unspecified isomorphism in Proposition 3.2 does not a priori enable such a control, but the one in Proposition 3.3 enables this easily. Let us state this refined regularity result.

Theorem 3.5. *Let s be a positive integer, let $\epsilon > 0$, and let $\mathcal{F} : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ be a growth function. Then there is a real number $M = O_{s,\epsilon,\mathcal{F}}(1)$ such that for any prime number $p \geq N_0(s, \epsilon, \mathcal{F})$ and any function $f : \mathbb{Z}_p \rightarrow [0, 1]$ there is a decomposition*

$$f = f_{\text{nil}} + f_{\text{sml}} + f_{\text{unf}}$$

with the following properties:

- (i) $f_{\text{nil}} : \mathbb{Z}_p \rightarrow [0, 1]$ is a p -periodic, $\mathcal{F}(M)$ -irrational nilsequence of degree at most s and complexity at most M , with underlying filtered nilmanifold of the form $Y_1 \times Y_2$ where Y_1 is a 2-step nilmanifold equipped with the lower central series, and Y_2 is a torus equipped with the maximal degree-2 filtration.
- (ii) $f_{\text{sml}} : \mathbb{Z}_p \rightarrow [-1, 1]$ satisfies $\|f_{\text{sml}}\|_2 \leq \epsilon$.
- (iii) $f_{\text{unf}} : \mathbb{Z}_p \rightarrow [-1, 1]$ satisfies $\|f_{\text{unf}}\|_{U^{s+1}} \leq 1/\mathcal{F}(M)$.
- (iv) $f_{\text{nil}} + f_{\text{sml}}$ takes values in $[0, 1]$.

We can now reduce the proof of Theorem 2.15 as follows.

Proposition 3.6. *Suppose that $(\mathbb{T}, \mathbb{R}_{\bullet(1)})$ models every 2-step nilmanifold with lower central series, and that $(\mathbb{T}, \mathbb{R}_{\bullet(2)})$ models every torus with maximal degree-2 filtration. Then Theorem 2.15 holds.*

This can be proved using Proposition 3.2, but to give a detailed proof it is convenient to use the more precise Proposition 3.3.

Proof. Let $X = (G/\Gamma, G_\bullet)$ be a filtered nilmanifold of degree at most 2, fix any basis \mathcal{X} on X , and let $\delta > 0$. We apply Proposition 3.3, and let $Y_1 = (H/\Gamma_H, H_{\bullet(1)}, \mathcal{X}_H)$ and $Y_2 = (\mathbb{T}^m, \mathbb{R}_{\bullet(2)}^m)$ be the resulting nilmanifolds and θ be the resulting isomorphism $H \times \mathbb{R}^m \rightarrow G$ satisfying conditions (i) and (ii') from Definition 3.1. By assumption, for every $\delta_1, \delta_2 > 0$ there exists a δ_1 -balanced map $\phi_{g_1} : (\mathbb{T}, \mathbb{R}_{\bullet(1)}) \rightarrow (H/\Gamma_H, H_{\bullet(1)}, \mathcal{X}_H)$ and a δ_2 -balanced map $\phi_{g_2} : (\mathbb{T}, \mathbb{R}_{\bullet(2)}) \rightarrow (\mathbb{T}^m, \mathbb{R}_{\bullet(2)}^m)$ (recall the notation ϕ_g from the paragraph before Definition 2.11). Letting $g_1 \times g_2$ denote the product map $(r_1, r_2) \mapsto (g_1(r_1), g_2(r_2))$, we now let $g : \mathbb{R}^2 \rightarrow G$ be the composition $\theta \circ (g_1 \times g_2)$. Noting that θ is a $(\Gamma_H \times \mathbb{Z}^m, \Gamma)$ -consistent polynomial map and that $g_1 \times g_2$ is a $(\mathbb{Z}^2, \Gamma_H \times \mathbb{Z}^m)$ -consistent

polynomial map, we have that g is a (\mathbb{Z}^2, Γ) -consistent map in $\text{poly}_0(H_{\bullet(1)} \times \mathbb{R}_{\bullet(2)}^m, G_{\bullet})$. It now suffices to show that δ_1, δ_2 can be chosen in terms of δ, X so that ϕ_g is δ -balanced.

Let $\bar{\theta} : Y_1 \times Y_2 \rightarrow X$ denote the homeomorphism induced by θ , and note that $\bar{\theta}$ is a bilipschitz map with constant $O_Q(1)$, where Q is the complexity bound on X (this can be checked using results such as [12, Lemma A.17]). Given a system $\Psi : \mathbb{Z}^D \rightarrow \mathbb{Z}^t$ of size at most $1/\delta$ and a function $F_0 : X^\Psi \rightarrow \mathbb{C}$ with $\|F_0\|_{\text{Lip}} \leq 1$, it follows that $F := F_0 \circ (\bar{\theta}^t)$ is an $O_{Q,\delta}(1)$ -Lipschitz function on $(Y_1 \times Y_2)^\Psi$. Note that $(Y_1 \times Y_2)^\Psi$ is isomorphic to $Y_1^\Psi \times Y_2^\Psi$. Viewing F as a function on the latter product space, we may approximate F within $\delta/4$ in the supremum norm by a finite sum $\sum_{i \in [M]} F_i$, where for each i we have $F_i : (y_1, y_2) \mapsto F_{i,1}(y_1)F_{i,2}(y_2)$ where $F_{i,j}$ is $O_{Q,\delta}(1)$ -Lipschitz on Y_j^Ψ for $j = 1, 2$. (To see that such an approximation exists, one can first use that for any $C > 0$ the set of functions f on $Y_1^\Psi \times Y_2^\Psi$ with $\|f\|_{\text{Lip}} \leq C$ is totally bounded in the supremum norm; this follows from the Arzelà-Ascoli theorem. In particular, for $C = O_{Q,\delta}(1) \geq \|F\|_{\text{Lip}}$, there is a finite $\delta/8$ -net for the set of f on $Y_1^\Psi \times Y_2^\Psi$ with $\|f\|_{\text{Lip}} \leq C$. By the Stone-Weierstrass theorem, each function in this net is within $\delta/8$ in the supremum norm from a function F_i of the claimed form; hence F is within $\delta/4$ of such a function.) Now if for $j = 1, 2$ we have $\left| \int_{(\mathbb{T}, \mathbb{R}_{\bullet(j)})^\Psi} F_{i,j} \circ \phi_{g_j}^t - \int_{Y_j^\Psi} F_{i,j} \right| \leq \delta_j$, then $\left| \int_{X_2^\Psi} F_i \circ (\phi_{g_1} \times \phi_{g_2})^t - \int_{(Y_1 \times Y_2)^\Psi} F_i \right| \leq \delta_1 \|F_{i,2}\|_\infty + \delta_2 \|F_{i,1}\|_\infty$. It follows that we can choose δ_1, δ_2 in terms of δ, Q to obtain that $\left| \int_{X_2^\Psi} F_0 \circ \phi_g^t - \int_{X^\Psi} F_0 \right| \leq \delta$, as required. \square

4. BALANCED MAPS FROM THE CIRCLE TO 2-STEP NILMANIFOLDS WITH LOWER CENTRAL SERIES

Our aim here is to establish the following result.

Proposition 4.1. *Let $Y = (H/\Gamma, H_{\bullet(1)})$ be a 2-step nilmanifold with lower central series, and let $X_1 = (\mathbb{T}, \mathbb{R}_{\bullet(1)})$. Then X_1 models Y .*

From previous work we already have the discrete version of this proposition in which X_1 is replaced with \mathbb{Z}_p ; more precisely, we have the following more general result.

Proposition 4.2 (Existence of a balanced periodic polynomial sequence).

Let $\delta > 0$, and let $(G/\Gamma, G_{\bullet}, \mathcal{X})$ be a filtered based nilmanifold of complexity at most m . Then there exists $C = C(m, \delta) > 0$ such that for every prime $p \geq C$ there exists $g \in \text{poly}_0(\mathbb{Z}, G_{\bullet})$ that is $(p\mathbb{Z}, \Gamma)$ -consistent and such that the map $\phi_g : \mathbb{Z}_p \rightarrow G/\Gamma$, $n \mapsto g(n)\Gamma$ is δ -balanced. Moreover, if $G_{\bullet} = G_{\bullet(1)}$, then g can be taken to be linear.

A polynomial $g \in \text{poly}_0(\mathbb{Z}, G_{\bullet})$ is said to be *linear* if it is of the form $g(n) = g_1^n$, for some $g_1 \in G$.

Proposition 4.2 is essentially [2, Proposition 6.1]. To obtain the additional linearity claim for g above, the main fact used is that for the lower-central series $G_{\bullet(1)}$ there are no non-trivial i -th level characters with $i > 1$ (see [2, Definition 4.2]). It is then a simple task to find g_1 so that $g(n) = g_1^n$ has the desired properties (see [2, §6]).

From the ‘discrete time’ result Proposition 4.2, we shall deduce the ‘continuous time’ result Proposition 4.1. This can be done using the quotient-integral formula. Let us illustrate this in the simplest setting, namely the case of equidistribution just for 1-parameter orbits: if $g_1^n \Gamma$ is equidistributed in G/Γ , then for every $g_0 \in G$ the sequence $g_0 g_1^n \Gamma$ is also equidistributed in G/Γ . In particular, for every Lipschitz function F on G/Γ with $\|F\|_{\text{Lip}(\mathcal{X})} \leq 1$ and every $r \in [0, 1/p)$ we have⁸

$$\mathbb{E}_{n \in \mathbb{Z}_p} F(g_1^{r p + n} \Gamma) \approx_\epsilon \int_X F \, d\mu_X.$$

The periodicity and linearity of g imply that $\theta \mapsto g_1^{\theta p} \Gamma$ is a well-defined (continuous) map $\mathbb{T} \rightarrow G/\Gamma$. Then, by the quotient integral formula [6, Theorem 1.5.2], we have

$$\begin{aligned} \int_{\mathbb{T}} F(g_1^{\theta p} \Gamma) \, d\mu_{\mathbb{T}}(\theta) &= \int_{r \in [0, 1/p)} \mathbb{E}_{n \in \mathbb{Z}_p} F(g_1^{(r+n/p)p} \Gamma) \, p \, d\mu_{\mathbb{T}}(r) \\ &\approx_\epsilon \int_{r \in [0, 1/p)} \int_X F \, d\mu_X \, p \, d\mu_{\mathbb{T}}(r) = \int_X F \, d\mu_X. \end{aligned}$$

Let us now prove the general case.

Proof of Proposition 4.1. Fix a basis on $Y = (H/\Gamma, H_{\bullet(1)})$ and let $\delta > 0$. Our task is to produce a δ -balanced map $X_1 \rightarrow Y$. Let $g : \mathbb{Z} \rightarrow H$ be the polynomial map given by Proposition 4.2, such that the induced map $\phi_g : \mathbb{Z}_p \rightarrow H/\Gamma$, $n \mapsto g(n)\Gamma$ is δ -balanced. From periodicity we have that $g(n) = \gamma^{n/p}$ for some $\gamma \in \Gamma$. Let $\Psi : \mathbb{Z}^D \rightarrow \mathbb{Z}^t$ be a system of integer linear forms of size at most $1/\delta$. Then $g^t(\Psi(n))\Gamma^\Psi$ is δ -equidistributed in Y^Ψ . Let $\phi : \mathbb{T} \rightarrow H/\Gamma$ be the circle flow interpolating the map ϕ_g , that is $\phi : \mathbb{T} \rightarrow H/\Gamma$, $\theta \mapsto \gamma^\theta \Gamma$. Our main claim is that

$$\left| \int_{\mathbb{T}^D} F^{\otimes t} \circ \phi^t(\Psi(x)) \, d\mu_{\mathbb{T}^D}(x) - \int_{Y^\Psi} F^{\otimes t} \right| \leq \delta.$$

Using the notation $\gamma^v = \begin{pmatrix} \gamma^{v(1)} \\ \vdots \\ \gamma^{v(t)} \end{pmatrix} \in G^t$ for $v \in \mathbb{R}^t$ (see Definition A.1), the integral on the left side here is written $\int_{\mathbb{T}^D} F^{\otimes t}(\gamma^{\Psi(x)} \Gamma^t) \, d\mu_{\mathbb{T}^D}(x)$. By the quotient integral formula, this equals

$$\int_{[0, 1/p)^D} \left(\mathbb{E}_{n \in \mathbb{Z}_p^D} F^{\otimes t}(\gamma^{\Psi(r) + \Psi(\frac{n}{p})} \Gamma^t) \right) p^D \cdot \mu_{\mathbb{T}^D}(r)$$

⁸For $a, b \in \mathbb{C}$ we write $a \approx_\epsilon b$ if and only if $|a - b| \leq \epsilon$.

For each $r \in [0, 1/p)^D$, the \mathbb{Z}_p^D -orbit $(\gamma^{\Psi(r)} \cdot \gamma^{\Psi(\frac{n}{p})})\Gamma^t$ is still δ -equidistributed in G^Ψ/Γ^Ψ , since the equidistribution property is not affected by multiplying by the constant $\gamma^{\Psi(r)} \in G^\Psi$. Thus for each such r we have

$$\left| \mathbb{E}_{n \in \mathbb{Z}_p^D} F^{\otimes t}(\gamma^{\Psi(r)} \cdot \gamma^{\Psi(\frac{n}{p})}\Gamma^t) - \int_{Y^\Psi} F^{\otimes t} \right| \leq \delta,$$

and the result follows. \square

Let us end this section by using Proposition 4.2 to establish the easy case of Theorem 2.1, as follows.

Proposition 4.3. *For every $\epsilon > 0$ there exists $C > 0$ such that the following holds. For every measurable function $f : X_2 \rightarrow [0, 1]$ and every prime $p \geq C$, there is a function $f' : \mathbb{Z}_p \rightarrow [0, 1]$ such that, for every system Ψ of linear forms of size at most $1/\epsilon$ and complexity at most 2, we have $|S_\Psi(f : X_2) - S_\Psi(f' : \mathbb{Z}_p)| \leq \epsilon$.*

Proof. We first claim that f can be assumed to be continuous with Lipschitz norm depending only on ϵ , more precisely there exists f_0 with $\|f_0\|_{\text{Lip}(X_2)} \leq C'(\epsilon)$ such that $|S_\Psi(f : X_2) - S_\Psi(f_0 : X_2)| \leq \epsilon/2$ for every Ψ of size at most $1/\epsilon$. We prove this with the following compactness argument.

Let $\mathcal{L}(\epsilon)$ denote the finite set of all systems of forms Ψ of size at most $1/\epsilon$. Let us say that a point $w \in [0, 1]^{\mathcal{L}(\epsilon)}$ is *achieved* if there is a continuous function $f_w : X_2 \rightarrow [0, 1]$ such that $S_\Psi(f : X_2) = w(\Psi)$ for every $\Psi \in \mathcal{L}(\epsilon)$. Let W denote the closure of the set of achieved points in $[0, 1]^{\mathcal{L}(\epsilon)}$ with respect to the ℓ^∞ norm. By compactness of W there exists a finite $(\epsilon^2/4)$ -net W_ϵ of achieved points. Each continuous function f_w for $w \in W_\epsilon$ can be assumed to have finite Lipschitz norm, since the Lipschitz functions on X_2 are dense, relative to the supremum norm, in the set of continuous functions (by the Stone-Weierstrass theorem). Let $C' = \max\{\|f_w\|_{\text{Lip}(X_2)} : w \in W_\epsilon\}$. Now, given the measurable function f , first by Lusin's theorem there is a continuous function $f_c : X_2 \rightarrow [0, 1]$ such that $\|f - f_c\|_{L^1(X_2)} \leq \epsilon^2/4$, and we therefore have $|S_\Psi(f : X_2) - S_\Psi(f_c : X_2)| \leq \epsilon/4$ for all $\Psi \in \mathcal{L}(\epsilon)$ (using the L^1 -continuity described in Remark 2.9). Then for some $w \in W_\epsilon$ we have $\|f_c - f_w\|_\infty \leq \epsilon^2/4$, and it follows that $|S_\Psi(f : X_2) - S_\Psi(f_w : X_2)| \leq \epsilon/2$ for all $\Psi \in \mathcal{L}(\epsilon)$. Relabelling f_w as f_0 , our claim follows.

Now, given that $\|f_0\|_{\text{Lip}(X_2)} \leq C'(\epsilon)$, we may apply Proposition 4.2 with X_2 and δ sufficiently small depending only on ϵ , so that the resulting map $\phi_g : \mathbb{Z}_p \rightarrow X_2$ satisfies $|S_\Psi(f_0 : X_2) - S_\Psi(f_0 \circ \phi_g : \mathbb{Z}_p)| \leq \epsilon/2$ for all $\Psi \in \mathcal{L}(\epsilon)$. We can then take $f' = f_0 \circ \phi_g$. \square

5. BALANCED MAPS BETWEEN TORI OF DEGREE 2

For each positive integer d let us denote by $Y(d)$ the d -dimensional torus with maximal degree 2 filtration, $Y(d) = (\mathbb{T}^d, \mathbb{R}_{\bullet(2)}^d)$. The main result of this section is the following.

Proposition 5.1. *For every positive integer d we have that $Y(1)$ models $Y(d)$.*

For each positive integer k , we define the following continuous homomorphism:

$$\phi_k : \mathbb{T} \rightarrow \mathbb{T}^d, \quad x \mapsto (x, kx, k^2x, \dots, k^{d-1}x).$$

We shall prove the proposition by showing that for every $\delta > 0$, for k sufficiently large the map ϕ_k is a δ -balanced map $Y(1) \rightarrow Y(d)$. As we shall eventually see in Lemma 5.3 below, the main property enabling this is that, for every $j \in [d-1]$, for every $r_0, r_1, \dots, r_{j-1} \in \mathbb{Z}$ and $r_j \in \mathbb{Z} \setminus \{0\}$, we have $r_0 + r_1k + \dots + r_jk^j \neq 0$ for k sufficiently large (there are of course other choices of ϕ_k with this property).

Given a system of forms $\Psi : \mathbb{Z}^D \rightarrow \mathbb{Z}^t$ of size at most $1/\delta$, viewing Ψ as a matrix (cf. Definition A.1), for each $i \in [D]$ let u_i denote the i -th column $\Psi(e_i) \in \mathbb{Z}^t$ of Ψ (where e_1, \dots, e_D is the standard basis of \mathbb{R}^D). Recall from Definition A.1 that $\Psi^{[2]}$ is the subgroup of \mathbb{Z}^t generated by the collection of vectors consisting of the u_i , the products $u_i u_j$, and the vector binomial coefficients $\binom{u_i}{2}$. Let $v_1, \dots, v_m \in \mathbb{Z}^t$ be a set of generators for $\Psi^{[2]}$. We then have $\Psi^{[2]} = \mathbb{Z}v_1 + \dots + \mathbb{Z}v_m$ and we can express the Leibman nilmanifolds for Ψ on $Y(1), Y(d)$ as follows:

$$\begin{aligned} Y(1)^\Psi &= \mathbb{T}v_1 + \dots + \mathbb{T}v_m \leq \mathbb{T}^t, \\ Y(d)^\Psi &= \mathbb{T}^d v_1 + \dots + \mathbb{T}^d v_m \leq (\mathbb{T}^d)^t, \end{aligned}$$

where $\mathbb{T}^d v = \left\{ \begin{pmatrix} \theta v(1) \\ \vdots \\ \theta v(t) \end{pmatrix} : \theta \in \mathbb{T}^d \right\}$.

Now ϕ_k^t maps $Y(1)^\Psi$ into $Y(d)^\Psi$ (as can be checked directly or by Proposition A.3), and we want to show that this map is δ -equidistributed if k is large enough.

By Definition 2.10, we have to show that for any function $F : Y(d)^\Psi \rightarrow \mathbb{C}$ with $\|F\|_{\text{Lip}(\mathbb{T}^{dt})} \leq 1$, we have $\left| \int_{Y(d)^\Psi} F \, d\mu_{Y(d)^\Psi} - \int_{Y(1)^\Psi} F \circ \phi_k^t \, d\mu_{Y(1)^\Psi} \right| \leq \delta$. We use the following result on Fourier approximations of Lipschitz functions [11, Lemma A.9].

Lemma 5.2. *Let \mathbb{T}^r be the standard r -dimensional torus, with metric induced by the ℓ^∞ norm $\|(x_1, \dots, x_r)\|_{\mathbb{T}^r} := \sup_{1 \leq j \leq r} \|x_j\|_{\mathbb{T}}$. Let X be a subset of \mathbb{T}^r , and let $f : X \rightarrow \mathbb{C}$ be a Lipschitz function. Then for every positive integer N there exist $J = O_r(N^r)$, $c_1, \dots, c_J = O(\|f\|_\infty)$, and $m_1, \dots, m_J \in \mathbb{Z}^r$ such that for all $x \in X$ we have*

$$f(x) = \sum_{j=1}^J c_j e(m_j \cdot x) + O_r \left(\frac{\|f\|_{\text{Lip}(\mathbb{T}^r)} \log N}{N} \right).$$

Furthermore, the values of m_1, \dots, m_J depend on r, N but are otherwise independent of f or X .

We apply this in our situation, with $X = Y(d)^\Psi$, $r = td$, $f = F$. Choosing N sufficiently large, denoting the obtained characters $x \mapsto e(m_j \cdot x)$ by e_{m_j} , and letting $F_0 = \sum_{j=1}^J c_j e_{m_j}$, we have by the lemma that $\|F - F_0\|_\infty \leq \delta/4$. It therefore suffices to show that

$$\left| \int_{Y(d)^\Psi} F_0 \, d\mu_{Y(d)^\Psi} - \int_{Y(1)^\Psi} F_0 \circ \phi_k^t \, d\mu_{Y(1)^\Psi} \right| \leq \delta/2. \quad (13)$$

We shall in fact prove this with upper bound equal to 0, by showing that for sufficiently large k we have for every $j \in [J]$ that

$$\int_{Y(d)^\Psi} e_{m_j}(y) \, d\mu_{Y(d)^\Psi}(y) = \int_{Y(1)^\Psi} e_{m_j}(x, kx, \dots, k^{d-1}x) \, d\mu_{Y(1)^\Psi}(x). \quad (14)$$

Since (14) clearly holds when e_{m_j} restricted to $Y(d)^\Psi$ is the principal character, and otherwise we have $\int_{Y(d)^\Psi} e_{m_j}(y) \, d\mu_{Y(d)^\Psi}(y) = 0$, it will suffice to prove the following result.

Lemma 5.3. *Let χ be a character on $(\mathbb{T}^d)^t$ such that the restriction $\chi|_{Y(d)^\Psi}$ is not the principal character on $Y(d)^\Psi$. Then for every sufficiently large positive integer k , we have*

$$\int_{Y(1)^\Psi} \chi(x, kx, \dots, k^{d-1}x) \, d\mu_{Y(1)^\Psi}(x) = 0.$$

Indeed, applying this above for each $j \in [J]$, we deduce (13) for some $k \in \mathbb{N}$, as desired.

Proof. Since $(\mathbb{T}^d)^t \cong \bigoplus_{i \in [d]} \mathbb{T}^t$, there are characters χ_1, \dots, χ_d on \mathbb{T}^t such that

$$\chi(x, kx, \dots, k^{d-1}x) = \chi_1(x) \chi_2(kx) \cdots \chi_d(k^{d-1}x).$$

Moreover, since $Y(d)^\Psi \cong \bigoplus_{i \in [d]} Y(1)^\Psi$, at least one χ_j restricts to a non-principal character on $Y(1)^\Psi$. Let $j \in [d]$ be the greatest index such that this is the case. Then for k large enough, the character on $Y(1)^\Psi$ sending x to $\chi(x) = \chi_1(x) \chi_2(kx) \cdots \chi_j(k^{j-1}x)$ is non-trivial, since its frequency is a non-zero vector. The result follows. \square

With Propositions 5.1, 4.1, and 3.6, the proof of Theorem 2.15 is now complete, and this implies Proposition 2.8 (by Proposition 2.16). This together with Proposition 4.3 then gives us finally the transference result Theorem 2.1, and Theorem 1.5 follows.

6. ON MODELLING NILMANIFOLDS OF HIGHER DEGREE – A COUNTEREXAMPLE

The main result of this section concerns possible generalizations of Theorem 1.5 for systems of higher complexity. For each positive integer s let X_s denote the following filtered torus of degree s :

$$X_s = (\mathbb{T}^s, \mathbb{R}_{\bullet(1)} \times \cdots \times \mathbb{R}_{\bullet(s)}).$$

One may believe at first that the natural generalization of Theorem 1.5 for complexity $s > 2$ should consist in replacing X_2 with X_s . This generalization would hold if it were true that X_s models every filtered nilmanifold X of finite complexity and degree at most s . However, this claim fails already for $s = 3$, as we show in this section. More precisely, we prove that X_3 fails to model the nilmanifold of degree 3 defined as follows.

Let H be the Heisenberg group $\begin{pmatrix} 1 & \mathbb{R} & \mathbb{R} \\ & 1 & \mathbb{R} \\ & & 1 \end{pmatrix} := \left\{ \begin{pmatrix} 1 & x_1 & x_3 \\ & 1 & x_2 \\ & & 1 \end{pmatrix} : x_i \in \mathbb{R} \right\}$, let $\Gamma = \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ & 1 & \mathbb{Z} \\ & & 1 \end{pmatrix}$, and let H_{\bullet} denote the degree-3 filtration on H with $H_{(2)} = \begin{pmatrix} 1 & 0 & \mathbb{R} \\ & 1 & \mathbb{R} \\ & & 1 \end{pmatrix}$ and $H_{(3)} = \begin{pmatrix} 1 & 0 & \mathbb{R} \\ & 1 & 0 \\ & & 1 \end{pmatrix}$. The fact that this is indeed a filtration is checked directly; in particular since $H_{(2)}$ is abelian (isomorphic to \mathbb{R}^2), we have $[H_{(2)}, H_{(2)}] \subseteq H_{(4)} = \{\text{id}_H\}$. Let H_3 denote the filtered nilmanifold $(H/\Gamma, H_{\bullet})$. Let $\tilde{\pi}$ denote the quotient homomorphism $H \rightarrow H/H_{(3)} \cong \mathbb{R}^2$, and let π denote the induced projection $H/\Gamma \rightarrow \mathbb{T}^2$, defined by $\pi(x\Gamma) = \tilde{\pi}(x) + \mathbb{Z}^2$. Note that $\tilde{\pi}$ is a (Γ, \mathbb{Z}^2) -consistent map in $\text{poly}_0(H_{\bullet}, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)})$.

The main result of this section is the following.

Proposition 6.1. *The filtered torus X_3 does not model the filtered nilmanifold H_3 .*

Before going into the details, let us convey the idea of the proof.

Suppose for a contradiction that, for every fixed $\epsilon > 0$, there existed an ϵ -balanced map $\phi_g : X_3 \rightarrow H_3$, induced by a (\mathbb{Z}^3, Γ) -consistent continuous map

$$g \in \text{poly}_0(\mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)} \times \mathbb{R}_{\bullet(3)}, H_{\bullet}).$$

The idea is that, if ϵ is sufficiently small, then the assumed equidistribution property would imply that ϕ_g yields a type of continuous ‘cross section’ $\mathbb{T}^2 \rightarrow H/\Gamma$ which cannot exist. We make this idea precise in two main steps. The first step consists in showing that composing ϕ_g with the projection π induces a polynomial map $\beta : X_2 \rightarrow X_2$ which, if ϵ is sufficiently small, must be a certain type of surjective homomorphism (this is the combination of lemmas 6.2 and 6.4). This gives us precise information on the form that the polynomial g itself must have. In the second step we use this information to show, essentially, that a restriction of such a map ϕ_g to a cross section of \mathbb{T}^2 in \mathbb{T}^3 already yields a contradiction, because the assumed form of g in fact does not allow it to be a consistent polynomial (this is made precise in Lemma 6.5).

Let us now turn to the details. Let α denote the projection $\mathbb{T}^3 \rightarrow \mathbb{T}^2$, $x \mapsto (x_1, x_2)$ and note that α is a polynomial map $X_3 \rightarrow X_2$. We first show that the map $\pi \circ \phi_g$ factors through α .

Lemma 6.2. *There is a polynomial map $\beta : X_2 \rightarrow X_2$ such that $\pi \circ \phi_g = \beta \circ \alpha$.*

We use the following fact, which is a straightforward consequence of the definitions.

Lemma 6.3. *If $g \in \text{poly}(G_\bullet, G'_\bullet)$ is (Γ, Γ') -consistent and Γ, Γ' are normal subgroups of G, G' respectively, then the induced map ϕ_g is polynomial from G/Γ to G'/Γ' with respect to the quotient filtrations on these groups.*⁹

Proof of Lemma 6.2. For any two fixed points $a, b \in \mathbb{R}$, let

$$f_{a,b} : \mathbb{R} \rightarrow \mathbb{R}^2, \quad x \mapsto \tilde{\pi}(g(a, b, x)).$$

Since $x \mapsto g(a, b, x)$ and $\tilde{\pi}$ are polynomial, we have $f_{a,b} \in \text{poly}(\mathbb{R}_{\bullet(3)}, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)})$. From the (\mathbb{Z}^3, Γ) -consistency of g , we deduce that $f_{a,b}$ is $(\mathbb{Z}, \mathbb{Z}^2)$ -consistent. By Lemma 6.3, the induced map $\phi = \phi_{f_{a,b}} : \mathbb{T} \rightarrow \mathbb{T}^2$ is itself a polynomial map from $(\mathbb{T}, \mathbb{R}_{\bullet(3)})$ to X_2 . We claim that this map is constant. To prove this we show that the compositions of ϕ with each of the two coordinate projections on X_2 are constant.

Indeed, consider first the composition of ϕ with the projection to the first coordinate $X_2 \rightarrow X_1$. This is a continuous polynomial map $\phi_1 : (\mathbb{T}, \mathbb{R}_{\bullet(3)}) \rightarrow X_1$, which implies that ϕ_1 is a polynomial map $\mathbb{T} \rightarrow \mathbb{T}$ of degree ≤ 1 , by Lemma 2.4. Moreover, as a polynomial map ϕ_1 must also conserve 3-dimensional cubes; this follows from Proposition A.3 applied to the system of forms Ψ_{U^3} corresponding to 3-cubes (recall from (6) that $\Psi_{U^3}(n_1, n_2, n_3, n_4) = (n_1 + v \cdot (n_2, n_3, n_4))_{v \in \{0,1\}^3}$). Now computing the Leibman group on $(\mathbb{T}, \mathbb{R}_{\bullet(3)})$ for this system, we find that $(\mathbb{T}, \mathbb{R}_{\bullet(3)})^{\Psi_{U^3}} \cong \mathbb{T}^{\{0,1\}^3}$. On the other hand, we have that $\phi_1 \circ c$ is a 3-cube over X_1 (i.e. lies in $(\mathbb{T}, \mathbb{R}_{\bullet(1)})^{\Psi_{U^3}}$) if and only if it is a function of $v \in \{0,1\}^3$ of the form $\phi_1 \circ c(v) = x + v \cdot h$, for some $x \in \mathbb{T}, h = (h_1, h_2, h_3) \in \mathbb{T}^3$. If ϕ_1 were non-constant on \mathbb{T} , then the function $\mathbb{T}^{\{0,1\}^3} \rightarrow \mathbb{T}^{\{0,1\}^3}$, $c \mapsto \phi_1 \circ c$ would have to map the 8-dimensional group $(\mathbb{T}, \mathbb{R}_{\bullet(3)})^{\Psi_{U^3}}$ to the 4-dimensional group $X_1^{\Psi_{U^3}}$, which is impossible. Hence ϕ_1 must be constant.

By a similar argument, the composition of ϕ with projection to the second coordinate on X_2 is shown to be constant, using the fact that the torus $(\mathbb{T}, \mathbb{R}_{\bullet(2)})^{\Psi_{U^3}}$ is 7-dimensional.

Now we define $\beta : X_2 \rightarrow X_2$, $(a + \mathbb{Z}, b + \mathbb{Z}) \mapsto \phi_{\tilde{\pi} \circ g}(a + \mathbb{Z}, b + \mathbb{Z}, x + \mathbb{Z})$, for a fixed $x + \mathbb{Z} \in \mathbb{T}$. It follows from the fact that $\phi_{f_{a,b}}$ is constant that β is independent of x , and we have $\pi \circ \phi_g = \beta \circ \alpha$. \square

⁹The i -th group in the quotient filtration of G_\bullet by Γ is $(G_{(i)} \cdot \Gamma)/\Gamma$.

Lemma 6.4. *If $\epsilon > 0$ is sufficiently small then $\beta(a, b) = (n_1a, n_2a + n_3b)$ where n_1, n_2, n_3 are integers and $n_1n_3 \neq 0$.*

Proof. First we claim that, as a polynomial map from X_2 to itself, β must be of the form

$$\beta(a, b) = (n_1a, n_2a + n_3b) \quad (15)$$

for integers n_1, n_2, n_3 . To see this, note first that composing β with coordinate projections and using Lemma 2.4, we have that $\beta(a, b) = (n_1a + n'_1b, n_2a + n_3b)$ for integers n_1, n'_1, n_2, n_3 . Moreover, since for each fixed a the map $b \mapsto n_1a + n'_1b$ is by assumption polynomial from $(\mathbb{T}, \mathbb{R}_{\bullet(2)})$ to $(\mathbb{T}, \mathbb{R}_{\bullet(1)})$, we must have $n'_1 = 0$, by considering cubes in an argument similar to the one in the previous proof. Thus (15) holds.

It remains to show that n_1, n_3 must both be non-zero.

If n_1 were zero, then the map β could not be ϵ -equidistributed for sufficiently small ϵ , and so ϕ_g could not be ϵ -balanced.

If n_3 were zero, then ϕ_g could still be ϵ -equidistributed, but it could not be ϵ -balanced for ϵ sufficiently small. Indeed, the map $\beta(a, b)$ would then depend only on a and would thus induce the polynomial map $\beta' : X_1 \rightarrow X_2$, $a \mapsto (n_1, n_2)a$. However, then β' cannot be ϵ -balanced for ϵ sufficiently small, as can be seen by considering for instance the system Ψ_{U^2} corresponding to 2-cubes (recall from (5) that $\Psi_{U^2}(n_1, n_2, n_3) = (n_1 + v \cdot (n_2, n_3))_{v \in \{0,1\}^2}$). Indeed, for any cube $c : \{0,1\}^2 \rightarrow \mathbb{T}$ in the Leibman group $(\mathbb{T}, \mathbb{R}_{\bullet(1)})^{\Psi_{U^2}}$, on one hand the linearity of β' implies that $\beta' \circ c$ lies in $(\mathbb{T}^2, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(1)})^{\Psi_{U^2}}$, this being the 6-dimensional subgroup of $(\mathbb{T}^2)^{\{0,1\}^2} \cong \mathbb{T}^8$ consisting of all the 2-cubes of degree 1, i.e. maps of the form $v \mapsto x + v_1h_1 + v_2h_2$ for some $x, h_1, h_2 \in \mathbb{T}^2$. On the other hand, the group $X_2^{\Psi_{U^2}} = (\mathbb{T}^2, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)})^{\Psi_{U^2}}$ is the greater, 7-dimensional subgroup of \mathbb{T}^8 consisting of maps $v \mapsto (x + v_1h_1 + v_2h_2, y_v)$, where x, h_1, h_2, y_v are seven independent parameters in \mathbb{T} . Consequently, for ϵ sufficiently small $\beta' \circ c$ cannot be ϵ -equidistributed in $X_2^{\Psi_{U^2}}$ as c ranges in the group $X_1^{\Psi_{U^2}}$. \square

Let us sum up the information on g that we have gathered so far.

The last two lemmas combined tell us that g must be a polynomial map from $(\mathbb{R}^3, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)} \times \mathbb{R}_{\bullet(3)})$ to (H, H_\bullet) of the form $g(a, b, c) = \begin{pmatrix} 1 & n_1a & r(a, b, c) \\ & 1 & n_2a + n_3b \\ & & 1 \end{pmatrix}$, where n_i are integers with n_1, n_3 non-zero, and where r is a real-valued polynomial in the real variables a, b, c such that g is (\mathbb{Z}^3, Γ) -consistent.

Letting $q(a, b) = r(a, b, 0)$, we deduce that $h(a, b) := \begin{pmatrix} 1 & n_1a & q(a, b) \\ & 1 & n_2a + n_3b \\ & & 1 \end{pmatrix}$ is a (\mathbb{Z}^2, Γ) -consistent continuous polynomial map from $(\mathbb{R}^2, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)})$ to (H, H_\bullet) . We shall now obtain a contradiction by examining the properties that q must satisfy.

Lemma 6.5. *Let $q : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a polynomial such that $h(a, b) := \begin{pmatrix} 1 & n_1 a & q(a, b) \\ & 1 & n_2 a + n_3 b \\ & & 1 \end{pmatrix}$ is a (\mathbb{Z}^2, Γ) -consistent polynomial map from $(\mathbb{R}^2, \mathbb{R}_{\bullet(1)} \times \mathbb{R}_{\bullet(2)})$ to (H, H_{\bullet}) . Then $n_1 n_3 = 0$.*

Proof. Given any real numbers a, b, k_1, k_2 , we have that $h(a, b)^{-1} h(a + k_1, b + k_2)$ equals

$$\begin{pmatrix} 1 & -n_1 a & -q(a, b) + n_1 a(n_2 a + n_3 b) \\ & 1 & -n_2 a - n_3 b \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & n_1(a + k_1) & q(a + k_1, b + k_2) \\ & 1 & n_2(a + k_1) + n_3(b + k_2) \\ & & 1 \end{pmatrix}.$$

The top-right entry in this matrix product equals

$$q(a + k_1, b + k_2) - n_1 a(n_2(a + k_1) + n_3(b + k_2)) - q(a, b) + n_1 a(n_2 a + n_3 b).$$

Simplifying this, we obtain the following 4-variable polynomial:

$$q'(a, b, k_1, k_2) = q(a + k_1, b + k_2) - q(a, b) - n_1 a(n_2 k_1 + n_3 k_2).$$

Now the assumed (\mathbb{Z}^2, Γ) -consistency implies that for every integer values of k_1, k_2 we have $q'(a, b, k_1, k_2) \in \mathbb{Z}$ for every $a, b \in \mathbb{R}$. We claim that this implies that $q'(a, b, k_1, k_2)$ is in fact independent of a and b .

To see this, let us split q' into the sum $q_0 + q_1$, where $q_0(a, b, k_1, k_2)$ consists of all the monomials of q' that involve at least one of a or b , and $q_1(k_1, k_2)$ consists of the monomials of q' involving only k_1 or k_2 . If q_0 were not the zero polynomial, then there would have to be integers m_1, m_2 such that $(a, b) \mapsto q_0(a, b, m_1, m_2)$ is not the zero polynomial. However, our assumption then implies that for every $a, b \in \mathbb{R}$ we have $q_1(m_1, m_2) + q_0(a, b, m_1, m_2) \in \mathbb{Z}$, which is impossible if $(a, b) \mapsto q_0(a, b, m_1, m_2)$ is not identically zero. Hence $q_0(a, b, k_1, k_2)$ must be identically zero and our claim follows.

We thus have a 2-variable real polynomial $q'(k_1, k_2)$ that satisfies

$$q'(k_1, k_2) = q(a + k_1, b + k_2) - q(a, b) - n_1 a(n_2 k_1 + n_3 k_2), \quad \forall a, b, k_1, k_2 \in \mathbb{R}. \quad (16)$$

Let us now consider the partial derivatives of q' around $(0, 0)$.

Noting that $q'(0, 0) = 0$ (by (16)), we have for every $a, b \in \mathbb{R}$ that

$$\partial_1 q'(0, 0) := \lim_{\epsilon \rightarrow 0} \frac{q'(\epsilon, 0)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{q(a + \epsilon, b) - q(a, b)}{\epsilon} - n_1 n_2 a = \partial_1 q(a, b) - n_1 n_2 a.$$

Letting c_1 denote the constant $\partial_1 q'(0, 0)$, we deduce that $\partial_1 q(a, b) = n_1 n_2 a + c_1$ for every $a, b \in \mathbb{R}$. Integrating with respect to a , this implies that for some 1-variable real polynomial t we have

$$q(a, b) = n_1 n_2 a^2 / 2 + c_1 a + t(b). \quad (17)$$

With respect to the second variable, we have for every $a, b \in \mathbb{R}$ that

$$\partial_2 q'(0, 0) := \lim_{\epsilon \rightarrow 0} \frac{q'(0, \epsilon)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{q(a, b + \epsilon) - q(a, b)}{\epsilon} - n_1 n_3 a = \partial_2 q(a, b) - n_1 n_3 a.$$

Denoting the constant $\partial_2 q'(0, 0)$ by c_2 , we have then $\partial_2 q(a, b) = n_1 n_3 a + c_2$ for all $a, b \in \mathbb{R}$. Hence for some 1-variable real polynomial t' we have

$$q(a, b) = n_1 n_3 ab + c_2 b + t'(a). \quad (18)$$

The equations (17) and (18) are consistent only if $n_1 n_3 ab$ is identically zero, that is only if $n_1 n_3 = 0$. \square

Fixing $\epsilon > 0$ sufficiently small, we have that Lemma 6.5 contradicts Lemma 6.4, and this completes the proof of Proposition 6.1.

7. FINAL REMARKS

In the last decade, a general approach has emerged in the study of very large combinatorial structures which proceeds by relating these structures to infinite continuous objects. A central example is the notion of limit objects for convergent sequences of graphs, objects which can be represented by symmetric two-variable measurable functions $W : [0, 1]^2 \rightarrow [0, 1]$ called *graphons* [18]. Analogous limit objects can be defined in the arithmetic setting for certain notions of convergence for sequences of functions on abelian groups [21, §6.2]. The concept in this paper of a continuous model for the convergent sequences $m_\Psi(\alpha, \mathbb{Z}_p)$ can be viewed as part of the general approach, but it differs from the study of limit objects in a significant way that we would like to emphasize here.

A useful definition of convergence for a sequence of functions (f_n) on abelian groups Z_n should guarantee in particular that certain averages of f_n with respect to linear configurations also converge. This is the case for instance in [21] for linear configurations of complexity 1. One of the central purposes of a limit object for such a sequence (f_n) , in this case a measurable function f on some compact abelian group, is that the average of f for each such configuration should be *equal* to the limit of the same averages for f_n over Z_n . In this sense the notion of limit object involves *exactness*. By contrast, the notion of a continuous model X requires only that one be able to *approximate* the average of f_n by the average of a function over X , with arbitrary prescribed accuracy, provided that n is sufficiently large. With this notion of a model, we lose the exactness, but we gain in that the model can be much simpler than a limit object. An example of this gain is the fact that in [21, Theorem 1] the limit object for systems of complexity 1 on groups \mathbb{Z}_p is a compact abelian group with a priori unbounded dimension, whereas in [1] it is shown that for modelling such systems the circle group suffices.

Let us mention that there is an alternative way to phrase most results in this paper, namely by working with nilspaces and morphisms between them. While this language

can simplify certain formulations and be clearer conceptually, there is presently less background literature on it than on filtered nilmanifolds and polynomial maps, so we leave the use of this language for future work.

Finally, note that the notion of modelling from Definition 2.13 yields a preorder on the set of filtered nilmanifolds, namely the preorder defined by $X \leq Y$ if and only if X models Y . Solving the following problem concerning this preorder would yield a generalization of Theorem 1.5.

Problem 7.1. *For each $s > 2$, find a filtered nilmanifold X_s of minimal dimension that is a lower bound, in the modelling preorder, for the set of filtered nilmanifolds of degree at most s .*

APPENDIX A. BACKGROUND NOTIONS AND RESULTS

By a filtration G_\bullet on a group G we mean a sequence of nested subgroups $G_{(0)} \supset G_{(1)} \supset \dots$ of G with $G_{(0)} = G_{(1)} = G$ and such that for every $i, j \geq 0$ the group of commutators $[G_{(i)}, G_{(j)}]$ is contained in $G_{(i+j)}$. We say that G_\bullet has degree at most d if $G_{(d+1)} = \{\text{id}_G\}$.

We shall use the ring structure on \mathbb{R}^t . With this structure, we can define binomial coefficients of vectors $v = \begin{pmatrix} v(1) \\ \vdots \\ v(t) \end{pmatrix} \in \mathbb{R}^t$, writing $\binom{v}{k}$ for the vector $\begin{pmatrix} \binom{v(1)}{k} \\ \vdots \\ \binom{v(t)}{k} \end{pmatrix} \in \mathbb{R}^t$, for k a non-negative integer.

Definition A.1 (The Leibman group). Let $\Psi = (\psi_1, \dots, \psi_t)$ be a collection of linear forms $\psi_1, \dots, \psi_t : \mathbb{Z}^D \rightarrow \mathbb{Z}$, let us view Ψ as a matrix in $\mathbb{Z}^{t \times D}$, and let v_1, \dots, v_D denote the columns of Ψ . For each $i \geq 1$, we denote by $\Psi^{[i]}$ the subgroup of \mathbb{Z}^t generated by the elements $\prod_{j \in [D]} \binom{v_j}{k_j}$ where $1 \leq \sum_{j \in [D]} k_j \leq i$. Given a filtered group (G, G_\bullet) , the *Leibman group* for Ψ on (G, G_\bullet) is the following normal subgroup of G^t :

$$(G, G_\bullet)^\Psi = \langle g^v : g \in G_{(i)}, v \in \Psi^{[i]}, i \geq 1 \rangle,$$

where $g^v := \begin{pmatrix} g^{v(1)} \\ \vdots \\ g^{v(t)} \end{pmatrix} \in G^t$.

This is essentially Leibman's original definition [17, §5.6], except that the filtration here is a general one (instead of the lower central series) and we are not assuming that the subgroup generated by the columns of Ψ contains $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ (unlike in [17, §5.2]).¹⁰

¹⁰In [13, Definition 3.1] the definition of the Leibman group involves real subspaces of \mathbb{R}^t rather than subgroups of \mathbb{Z}^t , and is restricted to connected and simply-connected Lie groups G . However, for such groups, that definition agrees with the one given here; this fact was already noted in [17, §5.9].

Recall the following general notion of polynomial maps between filtered groups (this is a special case of [14, Definition B.1]).

Definition A.2 (Polynomial maps). Let (G, G_\bullet) and (H, H_\bullet) be filtered groups. Given a map $g : H \rightarrow G$, and $h \in H$, we define the map $\partial_h g : H \rightarrow G$ by $\partial_h g(x) = g(x)^{-1}g(xh)$. We say that g is a *polynomial map* adapted to H_\bullet, G_\bullet if for every non-negative integers i_1, i_2, \dots, i_n and elements $h_j \in H_{(i_j)}$, $j \in [n]$, we have $\partial_{h_1} \partial_{h_2} \cdots \partial_{h_n} g(x) \in G_{(i_1 + \dots + i_n)}$ for all $x \in H$. The set of such maps is denoted $\text{poly}(H_\bullet, G_\bullet)$. We denote by $\text{poly}_0(H_\bullet, G_\bullet)$ the subset of $\text{poly}(H_\bullet, G_\bullet)$ consisting of those maps g satisfying $g(\text{id}_H) = \text{id}_G$.

We recall two central facts about these maps: firstly, that $\text{poly}(H_\bullet, G_\bullet)$ with pointwise product is a group (this is the Lazard-Leibman theorem [15, 16]); secondly, that polynomial maps enjoy the composition property, and so filtered groups with polynomial maps between them form a category (see [14, Corollaries B.4 and B.5]). Note that the same properties hold for $\text{poly}_0(H_\bullet, G_\bullet)$.

Note also that if G, H are abelian and we let H_\bullet be the lower central series and G_\bullet be the maximal degree- d filtration (cf. the paragraph before Theorem 1.5), then $g \in \text{poly}(H_\bullet, G_\bullet)$ if and only if g is polynomial of degree $\leq d$ in the sense of (8).

In Definition 2.11 and elsewhere we use the fact that polynomial maps with trivial constant term preserve Leibman groups. We justify this as follows.

Proposition A.3. *Let $(G, G_\bullet), (G', G'_\bullet)$ be filtered groups, let $g \in \text{poly}_0(G_\bullet, G'_\bullet)$, and let $\Psi : \mathbb{Z}^D \rightarrow \mathbb{Z}^t$ be a system of linear forms. Then g^Ψ maps $(G, G_\bullet)^\Psi$ into $(G', G'_\bullet)^\Psi$.*

Proof. Let us view Ψ as a matrix in $\mathbb{Z}^{t \times D}$ as in Definition A.1, with its i -th row identified with ψ_i as an element of \mathbb{Z}^D . Now note the following alternative characterization of the Leibman group:

$$(G, G_\bullet)^\Psi = \left\{ \begin{pmatrix} h(\psi_1) \\ \vdots \\ h(\psi_t) \end{pmatrix} : h \in \text{poly}_0(\mathbb{Z}^D, G_\bullet) \right\}. \quad (19)$$

This was already observed in [17, §5.7] in the case of the lower-central series $G_\bullet = G_{\bullet(1)}$. Let us outline a proof for general filtrations. To show that the right side of (19) is included in $(G, G_\bullet)^\Psi$, one can use the fact that every $h \in \text{poly}_0(\mathbb{Z}^D, G_\bullet)$ has a Taylor expansion (see [13, Lemma A.1]) with trivial 0-th coefficient. Substituting this expansion into $\begin{pmatrix} h(\psi_1) \\ \vdots \\ h(\psi_t) \end{pmatrix}$, we find that this is a product of elements of the form $g^{\prod_{j \in [D]} \binom{v_j}{k_j}}$ where $1 \leq \sum_{j \in [D]} k_j \leq i$ and $g \in G_{(i)}$, whence these elements are all in $(G, G_\bullet)^\Psi$, and the inclusion follows. The opposite inclusion follows from the fact that for each $g \in G_{(i)}$ and each v generating $\Psi^{[i]}$ (thus $v = \prod_{j \in [D]} \binom{\Psi(e_j)}{k_j}$) for the elements e_j of the standard

basis of \mathbb{R}^D), the element $g^v = \begin{pmatrix} g^{v(1)} \\ \vdots \\ g^{v(t)} \end{pmatrix}$ is of the form $\begin{pmatrix} h(\psi_1) \\ \vdots \\ h(\psi_t) \end{pmatrix}$ where h is the map $(n_1, \dots, n_D) \mapsto g^{\prod_{j \in [D]} \binom{n_j}{k_j}}$. By the same Taylor expansion result as above, the map h is in $\text{poly}_0(\mathbb{Z}^D, G_\bullet)$, so g^v is in the right side of (19). The inclusion is then fully deduced using the group property of $\text{poly}_0(\mathbb{Z}^D, G_\bullet)$.

To prove the main claim in the proposition, note that, by the composition property recalled above, we have for every $h \in \text{poly}_0(\mathbb{Z}^D, G_\bullet)$ that $g \circ h \in \text{poly}_0(\mathbb{Z}^D, G'_\bullet)$. Hence for every element $\begin{pmatrix} h(\psi_1) \\ \vdots \\ h(\psi_t) \end{pmatrix} \in (G, G_\bullet)^\Psi$ we have $g^t \begin{pmatrix} h(\psi_1) \\ \vdots \\ h(\psi_t) \end{pmatrix} = \begin{pmatrix} g \circ h(\psi_1) \\ \vdots \\ g \circ h(\psi_t) \end{pmatrix} \in (G', G'_\bullet)^\Psi$. \square

We now discuss nilmanifolds.

Definition A.4 (Filtered nilmanifold). A *filtered nilmanifold* $(G/\Gamma, G_\bullet)$ of degree at most d consists of a connected and simply-connected Lie group G , a lattice Γ in G (i.e. a discrete cocompact subgroup) and a filtration G_\bullet of degree at most d on G , with each group $G_{(i)}$ being a closed and connected subgroup of G such that $\Gamma_{(i)} := \Gamma \cap G_{(i)}$ is a lattice in $G_{(i)}$. If a Mal'cev basis¹¹ \mathcal{X} on G/Γ adapted to G_\bullet has been fixed, we refer to $(G/\Gamma, G_\bullet, \mathcal{X})$ as a *filtered based nilmanifold*. The *complexity* of such a nilmanifold is the least $C > 0$ such that the topological dimension of G , the degree of G_\bullet , and the rationality of \mathcal{X} are all at most C .

Remark A.5. In this paper we occasionally use the term *possibly disconnected filtered nilmanifold* to refer to a couple $(G/\Gamma, G_\bullet)$ where G is a simply-connected but not necessarily connected Lie group, where G_\bullet is a filtration of closed (simply-connected) subgroups $G_{(i)}$ of G , and where Γ is a lattice in G such that $\Gamma \cap G_{(i)}$ is a lattice in $G_{(i)}$. Let us emphasize that, by the term “filtered nilmanifold” without the additional term “possibly disconnected”, we always refer to the notion in Definition A.4.

When the filtration G_\bullet on G is clear from the context, we abbreviate the notation $(G, G_\bullet)^\Psi$ to G^Ψ .

Definition A.6 (Leibman nilmanifold). Let $X = (G/\Gamma, G_\bullet)$ be a filtered nilmanifold of finite degree, and let $\Psi : \mathbb{Z}^D \rightarrow \mathbb{Z}^t$ be a system of linear forms. The *Leibman nilmanifold* for Ψ on X , denoted X^Ψ , is the nilmanifold G^Ψ/Γ^Ψ .

For proofs that G^Ψ and Γ^Ψ have the required properties for G^Ψ/Γ^Ψ to be a nilmanifold, we refer the reader to [13, §3].

¹¹See [12, Definitions 2.1 and 2.4] for the definition of a Mal'cev basis and for the notion of quantitative rationality for such bases.

We now recall the definition of complexity of a system of forms, which was introduced by Gowers and Wolf (who called it “true complexity” to distinguish it from an earlier complexity notion introduced by Green and Tao) and studied in a series of papers beginning with [10].

Definition A.7 (Size and complexity of systems of linear forms). We say that a system $\Psi = (\psi_1, \dots, \psi_t)$ of linear forms $\psi_i : \mathbb{Z}^D \rightarrow \mathbb{Z}$ has *size at most* L if $D, t \leq L$ and the coefficients of each ψ_i have absolute value at most L . We say that Ψ has *complexity* s (on simple abelian groups) if for every $\eta > 0$ there exists $\lambda = \lambda(\Psi, \eta) > 0$ such that, for every prime p , for every $f, f' : \mathbb{Z}_p \rightarrow \mathbb{C}$ bounded by 1 and satisfying $\|f - f'\|_{U^{s+1}} \leq \lambda$, we have $|S_\Psi(f : \mathbb{Z}_p) - S_\Psi(f' : \mathbb{Z}_p)| \leq \eta$, and s is the least integer with this property.

There are equivalent formulations of this notion. The main one of these was originally conjectured by Gowers and Wolf in [10], and states that Ψ has complexity s if the powers ψ_i^{s+1} are linearly independent and s is the least integer with this property (here we view the forms ψ_i as linear polynomials with integer coefficients in $\mathbb{R}[x_1, \dots, x_D]$); see [9, 10], and also Theorem 7.1 and the remark at the end of Section 7 in [13]. This formulation makes it clear that every form in a system of finite complexity must be non-zero. This fact is relevant, for instance, in relation to the L^1 -continuity mentioned in Remark 2.9, as it implies the surjectivity of the coordinate projections mentioned in that remark.

Let us now turn to Theorem 2.14, which we restate here.

Theorem A.8. *For every $\delta > 0$ there exists $C > 0$ such that the following holds. For every prime $p \geq C$ and every function $f : \mathbb{Z}_p \rightarrow [0, 1]$, there is a filtered based nilmanifold $X = (G/\Gamma, G_\bullet, \mathcal{X})$ of degree at most 2 and complexity at most C , and a continuous function $F : G/\Gamma \rightarrow [0, 1]$ satisfying $\|F\|_{\text{Lip}(\mathcal{X})} \leq C$, such that for every system Ψ of integer linear forms of size at most $1/\delta$ and complexity at most 2, we have*

$$|S_\Psi(f : \mathbb{Z}_p) - S_\Psi(F : X)| \leq \delta. \quad (20)$$

The proof is essentially a combination of [2, Theorem 5.1] (a regularity result for the $U^d(\mathbb{Z}_p)$ norm) with [2, Theorem 4.1] (a counting result for balanced maps on \mathbb{Z}_p), but with a couple of additional observations.

Proof. Let $\eta > 0$ and let $\mathcal{F} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a growth function, both to be specified in terms of δ later, and let λ be as in Definition A.7. If $p \gg_{\eta, \mathcal{F}} 1$, we may apply the regularity result [2, Theorem 5.1] for the U^3 norm to f , to obtain a decomposition $f = f_{\text{nil}} + f_{\text{sml}} + f_{\text{unf}}$ and an integer $Q = O_{\eta, \delta, \mathcal{F}}(1)$ with the following properties:

- (i) There is a nilmanifold $X = (G/\Gamma, G_\bullet, \mathcal{X})$ of degree at most 2 and complexity at most Q , an $\mathcal{F}(Q)$ -irrational sequence $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ that is¹² p -periodic mod Γ , and $F : G/\Gamma \rightarrow \mathbb{C}$ with $\|F\|_{\text{Lip}(\mathcal{X})} \leq Q$, such that $f_{\text{nil}} = F(g(n)\Gamma)$. Note that we also have $g(0) = \text{id}_G$, indeed this follows from the factorization result [2, Proposition 5.2] as used in the proof of [2, Theorem 5.1].
- (ii) $\|f_{\text{sml}}\|_2 \leq \eta$,
- (iii) $\|f_{\text{unf}}\|_{U^3} \leq 1/\mathcal{F}(Q)$, and
- (iv) f_{nil} and $f_{\text{nil}} + f_{\text{sml}}$ take values in $[0, 1]$.

Furthermore, since f_{nil} is $[0, 1]$ -valued, we may assume that F is real-valued by taking real parts, and then by replacing it with $\max(\min(F, 1), 0)$ we may in fact assume that it is also $[0, 1]$ -valued; neither of these operations can increase $\|F\|_{\text{Lip}(\mathcal{X})}$.

Now let $\Psi = (\psi_1, \dots, \psi_t)$ be any system of linear forms $\psi_i : \mathbb{Z}^D \rightarrow \mathbb{Z}$ of size at most $1/\delta$ and complexity at most 2. Then by the choice of λ and a standard bound for $S_\Psi(f)$ in terms of $\|f\|_2$ (see [2, inequality (1)]), we have

$$|S_\Psi(f : \mathbb{Z}_p) - S_\Psi(f_{\text{nil}} : \mathbb{Z}_p)| = o_\delta(1) \quad \text{as } \eta \rightarrow 0 \text{ and } \mathcal{F}(Q) \rightarrow \infty. \quad (21)$$

We now deal with $S_\Psi(f_{\text{nil}} : \mathbb{Z}_p)$ using the counting result [2, Theorem 4.1]. The Lipschitz function that we use is $F^{\otimes t} : G^t/\Gamma^t \rightarrow \mathbb{C}$, $(x_1, \dots, x_t) \mapsto F(x_1) \cdots F(x_t)$. This satisfies $\|F^{\otimes t}\|_{\text{Lip}(\mathcal{X}^t)} = O_{Q,\delta}(1)$ by [2, Lemma A.4]. Applying [2, Theorem 4.1] to this function, with parameter $M = O_{\delta,\eta,\mathcal{F}}(1)$, we obtain

$$S_\Psi(f_{\text{nil}} : \mathbb{Z}_p) = \mathbb{E}_{n \in \mathbb{Z}_p^D} F^{\otimes t}(g^t(\Psi(n))\Gamma^t) = \int_{G^\Psi/\Gamma^\Psi} F^{\otimes t} + o_M(1)_{\mathcal{F}(Q) \rightarrow \infty}. \quad (22)$$

Combining (21) and (22), we now set $\eta = \eta(\delta)$ sufficiently small, and \mathcal{F} with sufficiently fast growth in terms of δ , to obtain (12), noting that $\int_{G^\Psi/\Gamma^\Psi} F^{\otimes t} = S_\Psi(F : X)$. \square

Acknowledgements. This work was supported by the ERC Consolidator Grant No. 617747. The authors also thank an anonymous referee for very useful remarks.

REFERENCES

- [1] P. Candela, O. Sisask, *On the asymptotic maximal density of a set avoiding solutions to linear equations modulo a prime*, Acta Math. Hungar. **132** (2011) (3), 223–243.
- [2] P. Candela, O. Sisask, *Convergence results for systems of linear forms on cyclic groups, and periodic nilsequences*, SIAM J. Discrete Math. **28** (2014) (2), 786–810.
- [3] E. Croot, *The minimal number of three-term arithmetic progressions modulo a prime converges to a limit*, Canad. Math. Bull. **51** (2008), no. 1, 47–56.

¹²In the terminology of this paper, g is a $(p\mathbb{Z}, \Gamma)$ -consistent sequence in $\text{poly}(\mathbb{Z}, G_\bullet)$ such that $g(n)\Gamma$ is c -balanced with $c = o(1)_{\mathcal{F}(Q) \rightarrow \infty}$. For the definition of irrationality, see [2, Definition 4.7].

- [4] E. Croot, V. Lev, *Open problems in additive combinatorics*, Additive Combinatorics, CRM Proceedings & Lecture Notes, vol. 43, 2007.
- [5] L. Corwin, F.P. Greenleaf, *Representations of nilpotent Lie groups and their applications, Part 1: Basic theory and examples*, Cambridge studies in advanced mathematics 18, C.U.P. 1990.
- [6] A. Deitmar, S. Echterhoff, *Principles of harmonic analysis* (Springer, 2009).
- [7] T. Eisner, T. Tao, *Large values of the Gowers-Host-Kra seminorms*, J. Anal. Math. **117** (2012), 133–186.
- [8] W.T. Gowers, *A new proof of Szemerédi’s theorem*, GAFA **11** (2001), 465–588.
- [9] W.T. Gowers, J. Wolf, *Linear functions and quadratic uniformity for functions on \mathbb{Z}_N* , J. Anal. Math. **115** (2011), 121–186.
- [10] W.T. Gowers, J. Wolf, *The true complexity of a system of linear equations*, Proc. Lond. Math. Soc. (3) **100** (2010), no. 1, 155–176.
- [11] B. Green, T. Tao, *Quadratic uniformity of the Möbius function*, Ann. Inst. Fourier (Grenoble) **58** (2008), no. 6, 1863–1935.
- [12] B. Green, T. Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Ann. of Math. (2) **175** (2012), no. 2, 465–540.
- [13] B. Green, T. Tao, *An arithmetic regularity lemma, an associated counting lemma, and applications*, An irregular mind, 261–334, Bolyai Soc. Math. Stud., 21, Janos Bolyai Math. Soc., Budapest, 2010.
- [14] B. Green, T. Tao, T. Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$ -norm*, Ann. of Math. **176** (2012), no. 2, 1231–1372.
- [15] M. Lazard, *Sur les groupes nilpotents et les anneaux de Lie*, Ann. Sci. Ecole Normale Sup. (3) **71** (1954), 101–190.
- [16] A. Leibman, *Polynomial sequences in groups*, Journal of Algebra **201** (1998), 189–206.
- [17] A. Leibman, *Orbit of the diagonal in the power of a nilmanifold*, Trans. Amer. Math. Soc. **362** (2010), no. 3, 1619–1658.
- [18] L. Lovász, B. Szegedy, *Limits of dense graph sequences*, J. Comb. Theory, Ser. B **96** (2006), 933–957.
- [19] A. L. Onishchik, E. B. Vinberg, *Foundations of Lie Theory and Lie transformation groups*, Springer 1997.
- [20] O. Sisask, *Combinatorial properties of large subsets of abelian groups*, Ph.D. Thesis, University of Bristol (2009).
- [21] B. Szegedy, *Limits of functions on groups*, preprint. arXiv:1502.07861.
- [22] B. Szegedy, *On higher order Fourier analysis*, preprint. arXiv:1203.2260.
- [23] T. Tao, V. Vu, *Additive combinatorics*, Cambridge University Press, 2006.

E-mail address: pablo.candela@uam.es

E-mail address: szegedyb@gmail.com